Superresolution images reconstructed from aliased images

Patrick Vandewalle, Sabine Süsstrunk and Martin Vetterli

LCAV - School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Switzerland

ABSTRACT

In this paper, we present a simple method to almost quadruple the spatial resolution of aliased images. From a set of four low resolution, undersampled and shifted images, a new image is constructed with almost twice the resolution in each dimension. The resulting image is aliasing-free. A small aliasing-free part of the frequency domain of the images is used to compute the exact subpixel shifts. When the relative image positions are known, a higher resolution image can be constructed using the Papoulis-Gerchberg algorithm. The proposed method is tested in a simulation where all simulation parameters are well controlled, and where the resulting image can be compared with its original. The algorithm is also applied to real, noisy images from a digital camera. Both experiments show very good results.

Keywords: Superresolution, subpixel motion estimation, aliased images

1. INTRODUCTION

Although digital cameras are appearing everywhere, a high resolution digital camera is still very expensive. The size of the CCD remains a limiting factor in the camera design. It would be much cheaper if a CCD with a small amount of pixels could be used several times to produce images of the same quality as a larger (and more expensive) CCD. This approach will be followed in this paper. We use a set of low resolution, undersampled images, which have been taken at slightly different positions, to reconstruct a high resolution image of the same scene.

Such a setup is typical in superresolution imaging. The concept was first introduced by Tsai and Huang.[1] They used a frequency domain approach for image registration and reconstruction of Landsat satellite image data. One of the main disadvantages of frequency domain methods is that they only allow for global motion of the entire scene. Another approach based on Bayesian and maximum likelihood (ML) methods is taken by Schultz et al.,[2] Elad and Feuer,[3] and Baker and Kanade.[4] Hendriks and van Vliet[5] make a comparison of different methods using cross-correlation and Taylor series. A good overview of the existing methods is given by Borman and Stevenson.[6]

The superresolution problem is most often divided into two tasks. First, image registration is required because the images are generally taken without knowing their relative positions. Some interesting frequency domain image registration algorithms are described by Kim and Su[7] and Stone et al.[8] Irani et al.[9] describe a spatial domain approach, which allows for multiple motions in the same image. Afterwards, the high resolution image can be reconstructed from the registered images. Cenker, Feichtinger, and Herrmann[10] describe different methods based on an alternating mapping, using the known information about the signal in a repetitive way. They also discuss possible criteria for quality/performance measurements. Strohmer[11] describes a reconstruction algorithm based on the solution of a linear system of equations, using two-dimensional trigonometric polynomials. The system of equations is solved using the conjugate gradient method. Keren et al.[12] and Hardie et al.[13] describe methods which start from an initial guess of the high resolution image. Then they improve this image by simulating the imaging process and minimizing the difference between the observed and simulated low resolution images.

Further author information: (Send correspondence to Patrick Vandewalle)
E-mail: {Patrick.Vandewalle, Sabine.Susstrunk, Martin.Vetterli}@epfl.ch, Telephone: +41 21 6935125, Fax: +41 21 6934312

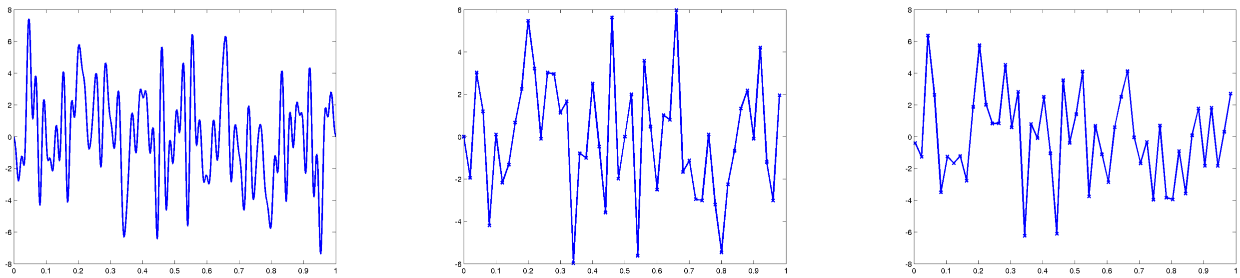
The images received by the CCD sensor of a digital camera have their bandwidth limited by the optical system. If the sampling density satisfies the Nyquist criterium, perfect reconstruction conditions are satisfied and the bandlimited image can be reconstructed from its sampled image points. On the other hand, if the sampling density is too low to satisfy the Nyquist criterium (or equivalently, if the cutoff frequency of the optical system is too high), the resulting image will be aliased and perfect reconstruction is impossible. More information is therefore needed. Our algorithm combines the information from four subpixel shifted low resolution images that are undersampled by a factor less than two in both dimensions. From these aliased images, we reconstruct an aliasing-free image with almost double resolution in both rows and columns.

This paper is structured as follows. In Section 2, we will discuss the algorithm for computing the subpixel shifts between the images. After image registration, a higher resolution image can be reconstructed (Section 3). Some results on artificially created and real images will then be shown in Section 4. In Section 5, results and possibilities for future work are discussed, and Section 6 concludes the article.

2. REGISTRATION

2.1. Shift estimation in one dimension

The registration algorithm is first developed for one-dimensional signals. $s(t)$ is a bandlimited continuous signal with maximal frequency f_{max}. This signal is sampled at times $t_0, t_0 + T, t_0 + 2T, ...t_0 + NT$ with a sampling frequency f_s, taking $f_s > f_{max}$. This results in a sampled signal $s_1[n], n = 0, 1, ..., N$. We now sample $s(t)$ again, at the same sampling frequency f_s, but at times $t_0 + \delta, t_0 + T + \delta, t_0 + 2T + \delta, ..., t_0 + NT + \delta$, which are shifted by an unknown shift δ compared to the first set of sampling times. We call this sampled signal $s_2[n], n = 0, 1, ..., N$. Figure 1 illustrates this concept.



(a) (b) (c)

Figure 1. (a) Original signal $s(t)$. (b) and (c) Sampled and aliased signals $s_1[n]$ and $s_2[n]$.

If f_s satisfies the Nyquist criterium, $f_s > 2f_{max}$, $s(t)$ can be perfectly reconstructed from $s_1[n]$ or $s_2[n]$ separately, and the shift δ - although not needed because $s(t)$ has already been reconstructed - can be derived from the two reconstructions using a correlation operator. One signal is sufficient to make a perfect signal reconstruction.

If we assume f_s has a value $f_{max} < f_s < 2f_{max}$, we know from Shannon sampling theory that a perfect reconstruction of $s(t)$ from $s_1[n]$ or $s_2[n]$ is not possible, because $s_1[n]$ and $s_2[n]$ will be aliased. Direct computation of δ using a correlation operator is impossible, because the aliasing effect causes the two signals to be different. But, because $f_{max} < f_s$, $s_1[n]$ and $s_2[n]$ also have a part that is aliasing-free, namely for the frequencies $|f| < f_s - f_{max}$. It is then possible to apply a lowpass filter to $s_1[n]$ and $s_2[n]$ with cutoff frequency $f_s - f_{max}$, which results in two identical, aliasing-free signals $s_{1,low}[n]$ and $s_{2,low}[n]$ (Figure 2). If we suppose $s(t)$ has non-zero energy in the frequency band $-f_s + f_{max} < f < f_s - f_{max}$, δ can be computed from $s_{1,low}[n]$ and $s_{2,low}[n]$ using a correlation operation.

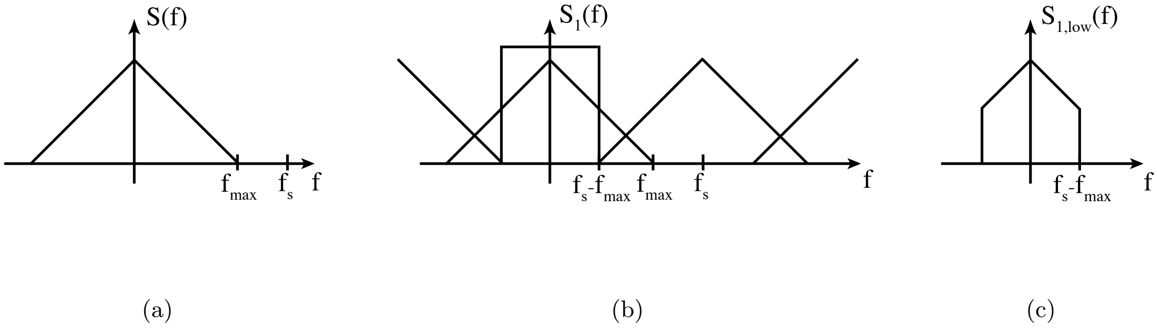


Figure 2. (a) Original signal $s(t)$. (b) Sampled and aliased signal $s_1[n]$ with the appropriate lowpass filter indicated. (c) Lowpass filtered signal $s_{1,low}[n]$

2.2. Motion estimation in 2D

This method can now easily be generalized to two-dimensional signals by performing all operations on the two signal dimensions (rows and columns) separately. We start from a 2D signal $r(x, y)$, which has maximal horizontal frequency $f_{max,h}$ and vertical frequency $f_{max,v}$. Sampling at

$$\begin{bmatrix} (x_0, y_0) & (x_0+T_x, y_0) & (x_0+2T_x, y_0) & \cdots & (x_0+MT_x, y_0) \\ (x_0, y_0+T_y) & (x_0+T_x, y_0+T_y) & (x_0+2T_x, y_0+T_y) & \cdots & (x_0+MT_x, y_0+T_y) \\ (x_0, y_0+2T_y) & (x_0+T_x, y_0+2T_y) & (x_0+2T_x, y_0+2T_y) & \cdots & (x_0+MT_x, y_0+2T_y) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (x_0, y_0+NT_y) & (x_0+T_x, y_0+NT_y) & (x_0+2T_x, y_0+NT_y) & \cdots & (x_0+MT_x, y_0+NT_y) \end{bmatrix}$$

for $r_1[k, l]$ and at

$$\begin{bmatrix} (x_0+\delta_x, y_0+\delta_y) & (x_0+T_x+\delta_x, y_0+\delta_y) & \cdots & (x_0+MT_x+\delta_x, y_0+\delta_y) \\ (x_0+\delta_x, y_0+T_y+\delta_y) & (x_0+T_x+\delta_x, y_0+T_y+\delta_y) & \cdots & (x_0+MT_x+\delta_x, y_0+T_y+\delta_y) \\ (x_0+\delta_x, y_0+2T_y+\delta_y) & (x_0+T_x+\delta_x, y_0+2T_y+\delta_y) & \cdots & (x_0+MT_x+\delta_x, y_0+2T_y+\delta_y) \\ \vdots & \vdots & \ddots & \vdots \\ (x_0+\delta_x, y_0+NT_y+\delta_y) & (x_0+T_x+\delta_x, y_0+NT_y+\delta_y) & \cdots & (x_0+MT_x+\delta_x, y_0+NT_y+\delta_y) \end{bmatrix}$$

for $r_2[k, l]$, the signal is sampled twice, with horizontal and vertical sampling frequencies $f_{max,h} < f_{s,h} < 2f_{max,h}$ and $f_{max,v} < f_{s,v} < 2f_{max,v}$. T_x and T_y are the horizontal and vertical sampling periods, respectively.

This results again in two aliased signals, which can then be lowpass filtered to $r_{1,low}[k, l]$ and $r_{2,low}[k, l]$ (with filters having cutoff frequencies $f_{c,h} = f_{s,h} - f_{max,h}$ and $f_{c,v} = f_{s,v} - f_{max,v}$). From these two filtered signals, $\delta = (\delta_x, \delta_y)$ can be derived using a two-dimensional correlation operator. A similar method for 2D motion estimation was used by Kim and Su[7] and Stone et al.[8]

3. RECONSTRUCTION

We use the Papoulis Gerchberg algorithm[14,15] for the construction of the high resolution image. This algorithm is a special case of the Projection Onto Convex Sets (POCS) method. We assume that the image belongs to two convex, linear sets. Through repeated projections, the algorithm converges iteratively to the desired image at the intersection of the two sets.

The values at unknown pixel positions are set to zero during initialization. At each iteration, two projections are performed. First the image is projected onto the set of bandlimited signals by setting the frequency values to zero for frequencies larger than the maximal allowed frequency f_{max}. Next, the image is projected onto the set of unknown samples by assigning the correct values again to the known samples.

This algorithm converges to the desired solution if the number of known samples is larger than the number of unknown Fourier coefficients. Sampling at a frequency $f_s > f_{max}$ results in N^2 sample values per image, for $(2K)^2 = 4K^2$ Fourier coefficients ($N > K$). Therefore, a total of $\lceil 4K^2/N^2 \rceil$ images will be needed to reconstruct a good high resolution image. If we assume that $K < N$ but close to N, four images will be needed for reconstruction to have maximal gain in resolution. The effective bandwidth of the resulting (reconstructed) image is $(f_{max} - \epsilon)^2$, compared to $(f_{max}/2)^2$ for the original, aliased images.

4. RESULTS

In this section, some experiments with the described method and their results will be discussed. In Section 4.1, we will describe a simulation where four low resolution images were created from an original high resolution image. It is then possible to verify if the reconstruction conditions are satisfied. The results can be compared with the ground truth image, which is not available in an experimental setup. In Section 4.2, an experiment is described where a high resolution image is created from four images, captured by a digital camera, without exact knowledge about the parameters.

4.1. Simulation

First, we test the algorithm on an artificial set of images in order to have as much control as possible over the different parameters. We start from an original image (256x256 pixels), which we consider as the equivalent of the real continuous space. This image is then filtered using an ideal lowpass filter with cutoff frequency $f_c = 0.125 f_{s,original} - \epsilon$, with $f_{s,original}$ the original sampling frequency of the high resolution image, and $\epsilon = 0.005 f_{s,original}$. Next, the filtered image is downsampled by a factor eight, keeping only every 8-th sample. Note that this is different from a normal lower-resolution image, where a set of pixel values from the high resolution image would normally be averaged to form a pixel of the low resolution image. Four downsampled images are derived like this, by shifting the set of selected pixels each time by a random number of pixels (0-7) in both horizontal and vertical directions.

This provides us exactly with the setup described above. We have four low resolution images (32x32 pixels), which are sampled at frequency f_s, with $f_s = f_{max} + \epsilon$ ($f_{max} = f_c$ is the maximal frequency in the image). We can then apply our algorithm to these low resolution images, compute the relative shifts between them and construct a high resolution image (see Figure 3). The reconstructed image is a very good approximation of the original (MSE=7.46e-6).

4.2. Experiment

The algorithm is also applied to a set of four images that were taken with a black and white digital camera (Figure 4). The camera was moved by a small amount for the different images, but the scene didn't change. Thus, we have only global motion. All four images have aliasing artifacts in the high frequency region (before blurring starts to occur). The extra difficulty in this case is that the sampling frequency and the maximal signal frequency are unknown, as is usually the case in realistic settings. As aliasing is only noticeable for relatively high frequencies, we assume the sampling frequency f_s can be approximated as $f_s = 1.1 f_{max}$, meaning that 10% of the frequency content of the signal is aliasing-free.

We apply the subpixel shift estimator to the four images, and we find relative shifts of (0.4,-0.3), (0.2,-0.2), (0.8,0.1) pixels for the second, third and fourth image, respectively, compared to the first image. Using these shift estimates, we can construct the high resolution image using the Papoulis-Gerchberg algorithm. Figure 5 shows that the aliasing has been removed from the images and a higher resolution image has been constructed.


(a) (b) (c) (d)

Figure 3. (a) Original image (256x256). (b) Lowpass filtered image (reconstruction target). (c) One of the four low resolution images to start from (32x32). (d) Reconstructed high resolution image. The MSE between images b and d is equal to 7.46e-6.

5. DISCUSSION AND FUTURE WORK

The price of CCD image sensors decreased sharply in the past few years. Cheap CCDs appear in many products: mobile phones, credit card sized digital cameras, surveillance equipment, etc. However, the images they produce often have very poor quality. They are blurred and suffer from aliasing artifacts. Therefore, it would be very interesting to produce high quality images from a large number of these lower quality images. The method described above almost doubles the maximal signal frequency that is atteignable in both rows and columns, re-

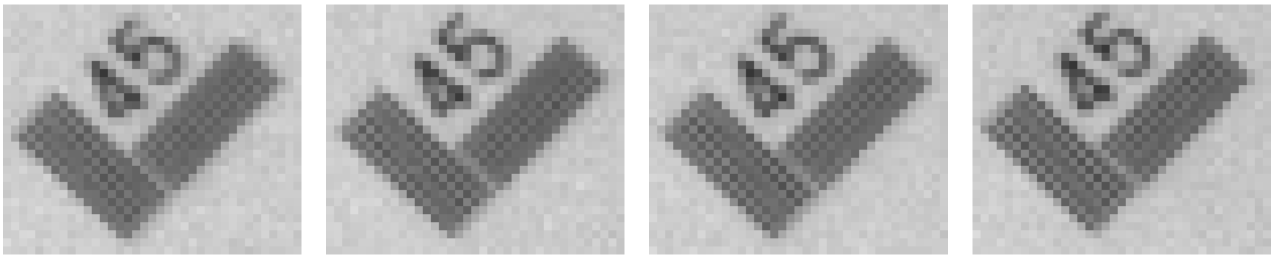
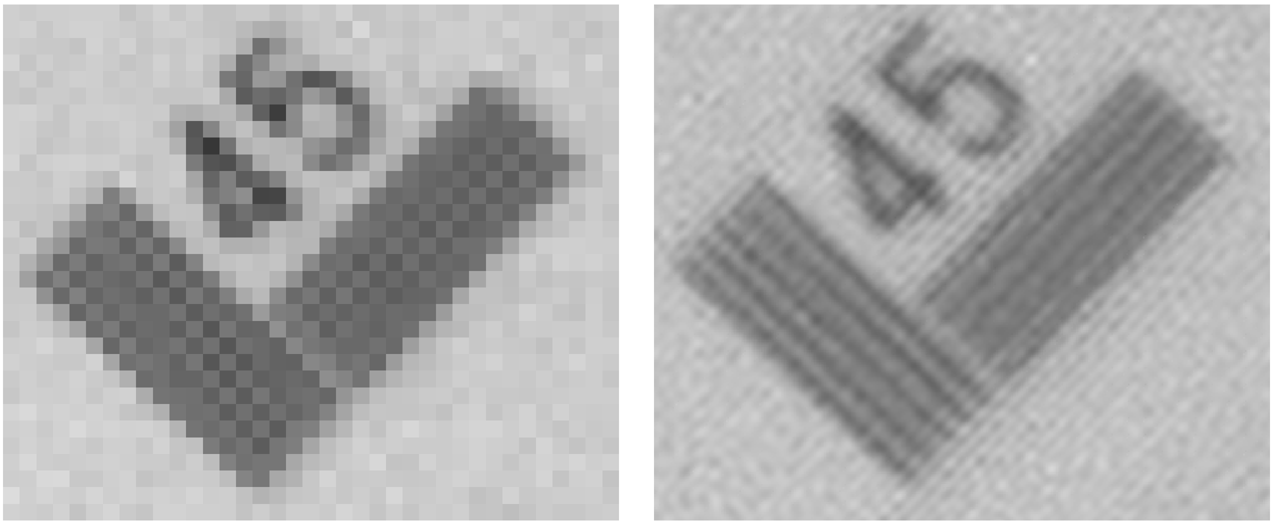


Figure 4. Four original (aliased) images taken with a black and white digital camera. The contrast in the images has been increased for illustration purposes.



(a) (b)

Figure 5. (a) One of the original images. (b) Resulting high resolution, aliasing-free image. The contrast in the images has been increased for illustration purposes.

moving aliasing that results from the typical sampling process. The whole motion estimation algorithm described above depends on the presence of a (small) aliasing-free part in the sampled signal. Because this aliasing-free part is exactly the lowest frequency part, the method is relatively insensitive to noise, which is typically a high frequency phenomenon.

The practical applicability of this method depends on the modulation transfer function (MTF) of the camera's optical system. In general, digital cameras are designed such that the optical system blocks frequencies that would cause aliasing. An optical system is chosen with an MTF which is zero at frequencies larger than the Nyquist frequency (given by the CCD sampling rate). Aliasing effects are avoided, at the expense of a blurring of the image. For images taken with these types of cameras, our method is not effective. It only works for cameras where frequencies larger than the Nyquist frequency pass through the optical system onto the CCD. On the other hand, most color digital cameras use a CFA filter array. This type of array usually has twice as many sensors for green as for red or blue. The optical system is designed to eliminate aliasing in the green channel of the image. Because there are only half as many red and blue pixels, aliasing is possible in those channels. They are subsampled by a factor smaller than two in both dimensions, so our method is perfectly applicable and could produce very good results.

Future work will include the analysis and development of a new reconstruction method which makes better use of the regularity in the set of samples. In fact, the sample positions are not totally random. They form a set of periodic nonuniform samples, for which more efficient reconstruction methods can be developed.

In the current simulations, we assume the signal is sampled using a Dirac sampling proces. In reality, of course, we need to use a point spread function to model the sampling process. This causes the image to be blurred compared to an image sampled with Dirac sampling. It would be very interesting to include a realistic point spread function in our model and analyze its effects on the resulting high resolution image.

6. CONCLUSIONS

In this paper, we presented a new method to create a high resolution image from four undersampled, aliased images. The new image has almost double resolution in both dimensions. The method is based on an exact computation of the relative motion between the images from an aliasing-free part in the images. The high resolution image is then reconstructed using the Papoulis-Gerchberg algorithm.

Simulations and experimental results show that this method also works in practice. In the simulations, the ground truth image was reconstructed with very small error. The experiment shows an image from which all aliasing is accurately removed.

ACKNOWLEDGMENTS

The work presented in this paper was supported by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322. We would also like to thank Urs Schmid from Leica Microsystems AG for providing the digital pictures.

REFERENCES

1. T. S. Huang, ed., *Advances in Computer Vision and Image Processing*, vol. 1, ch. 7, pp. 317–339. JAI Press, 1984.
2. R. R. Schultz, L. Meng, and R. L. Stevenson, "Subpixel motion estimation for super-resolution image sequence enhancement," *Journal of Visual Communication and Image Representation* **9**, pp. 38–50, March 1998.
3. M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," *IEEE Transactions on Image Processing* **6**, pp. 1646–1658, December 1997.
4. S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**, pp. 1167–1183, September 2002.
5. C. L. Hendriks and L. van Vliet, "Improving resolution to reduce aliasing in an undersampled image sequence," in *Proceedings SPIE Electronic Imaging 2000 Conference San Jose*, **3965**, pp. 214–222, January 2000.
6. S. Borman and R. Stevenson, "Spatial resolution enhancement of low-resolution image sequences - a comprehensive review with directions for future research," tech. rep., University of Notre Dame, 1998.
7. S. P. Kim and W.-Y. Su, "Subpixel accuracy image registration by spectrum cancellation," in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, **5**, pp. 153–156, April 1993.
8. H. S. Stone, M. T. Orchard, E.-C. Chang, and S. A. Martucci, "A fast direct fourier-based algorithm for subpixel registration of images," *IEEE Transactions on Geoscience and Remote Sensing* **39**, pp. 2235–2243, October 2001.
9. M. Irani, B. Rousso, and S. Peleg, "Computing occluding and transparent motions," *International Journal of Computer Vision* **12**, pp. 5–16, February 1994.

10. C. Cenker, H. G. Feichtinger, and M. Herrmann, "Iterative algorithms in irregular sampling - a first comparison of methods," in *IEEE Conference Proceedings of the Tenth Annual International Phoenix Conference on Computers and Communications (IPCCC)*, pp. 483–489, March 1991.
11. T. Strohmer, "Computationally attractive reconstruction of bandlimited images from irregular samples," *IEEE Transactions on Image Processing* **6**, pp. 540–548, April 1997.
12. D. Keren, S. Peleg, and R. Brada, "Image sequence enhancement using sub-pixel displacements," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pp. 742–746, June 1988.
13. R. C. Hardie, T. R. Tuinstra, J. Bognar, K. J. Barnard, and E. Armstrong, "High resolution image reconstruction from digital video with global and non-global scene motion," in *IEEE Proceedings International Conference on Image Processing*, **1**, pp. 26–29, October 1997.
14. A. Papoulis, "A new algorithm in spectral analysis and band-limited extrapolation," *IEEE Transactions on Circuits and Systems* **22**(9), pp. 735–742, 1975.
15. R. W. Gerchberg, "Super-resolution through error energy reduction," *Optica Acta* **21**(9), pp. 709–720, 1974.