Single-Image Digital Super-Resolution A Revised Gerchberg-Papoulis Algorithm

Eran Gur, *Member, IAENG* and Zeev Zalevsky

***Abstract*—The subject of extracting high-resolution data from low-resolution images is one of the most important digital processing applications in recent years, attracting much research. In this work the authors show how to improve the resolution of an image when a small part of the image is given in high-resolution. To obtain this result the authors use an iterative procedure imposing the low frequencies complete data of the original low-resolution image and the high-resolution data present only in a fraction of the image. The procedure is based on the Gerchberg-Papoulis algorithm and contains dynamic properties, not present in the original scheme. The result is a clearer image, with higher correlation to the required high-resolution image. The authors show the use of such a procedure on a rosette image and on a facial image to demonstrate the higher frequencies obtained and on a text sample to show improvement in textual understanding.**

***Index Terms*—Image Processing, Signal reconstruction, Super- resolution.**

I. Introduction

In recent years the subject of super resolution (SR) has become very popular. SR refers to recovering high resolution data from images that due to mis-focus, compression or other forms of distortion have lost the data that were originally embedded in the higher frequencies of the image, and hence are now given as low resolution images. The methods to overcome this problem of data loss, and generate SR, are quite versatile. In some cases the method is to obtain data concerning the blurring function and use an inverse filter to reconstruct the high-resolution image [1,2]. Unfortunately, two main problems limit this approach. The first, usually it is impossible to identify the exact blurring function since it is a result of stochastic noise and thus only its statistical properties are known.

The second problem is that, even if the blurring function is known making an inverse filter might not be practical (e.g. if the original blurring filter has zeros the inverse filter must have singular values to obtain exact restoration). Other methods use large databases; they are divided into two groups. In the first group [3,4] one takes a large amount of different test-images present both in a low resolution version and a high resolution version, and attempt to find the blurring procedure that will yield the best results with respect to all images. There are two problems with this approach, no two pictures are identical and therefore we cannot be sure that the inverse blurring procedure found will be applicable for the required new image, and, usually the blurring procedure varies from one test-image to another and thus the "anti-blurring" filter will be an average of many filters, and not an exact filter.

The second group of SR using large databases assumes one has many low-resolution pictures of the required subject [5,6]. Since in every picture a different portion of the high-resolution data is missing it is possible to extract some high-resolution data from these images to obtain a single high-resolution image. The main drawback of these methods is the large database required in order to increase the resolution of a single image.

In our work we suggest a novel approach assuming only one image is given – the required image. Suppose one has to scan an image with high resolution, it is time and memory consuming. However, if only a small portion of the image will be scanned at peak resolution and most of the data scanned with lower accuracy the process itself will become much faster and the storage capacity required to store the images will become much smaller (e.g. with CT scanning if the patient has to spend less time on the scanning device his/her inconvenience is reduced). This is the exact principle used in this work. We assume only a small portion of the required image is given in high-resolution and use this data (assuming it has similar statistical properties to the neighboring low-resolution portion of the image) to increase the resolution of the entire image. To do this we use an iterative procedure relying on two initial assumptions: in one small portion of the image we have all the high-resolution data, and the entire image contains all the low frequencies of the original high-resolution image. In the following sections we explain the procedure and show some test cases supporting this approach.

II. Iterative single image Super-Resolution

A. Concept

When using a single image we need to know the limits of our data. In our case we know for certain that the low frequencies exist in the low-resolution image as they would exist in the high-resolution image (we assume the blurring function has a relatively sharp frequency response, and thus the lower frequencies are not distorted). Thus we can impose a frequency-domain restriction on the image. We also know that a certain portion of the image is presented in high-resolution so we can impose an object-domain restriction on the image. These two restrictions allow us to

Manuscript submitted August 30, 2007

Eran Gur is with Shenkar College of Engineering & Design, Ramat Gan, 52526, Israel (corresponding author, phone: +972-3-6005367; fax: +972-3-6005370; e-mail: erangur@mail.shenkar.ac.il).

Zeev Zalevsky is with Bar Ilan University, Ramat Gan, 52900, Israel (e-mail: zalevsz@macs.biu.ac.il).

bounce back and forth from the object-domain to the frequency-domain, with a procedure similar to the one used for phase retrieval, as shown in the following subsection.

B. Review: Iterative Phase Retrieval

A well-known problem is to determine the phase of a phase only object plane filter that will produce a required intensity distribution in the Fourier domain. In their paper [7], Gerchberg and Saxton suggested an iterative procedure to do just that. It has been proven [8] that this procedure converges, in the sense that the MSE monotonically decreases as the number of iterations increases.

The concept is quite simple: We begin with an arbitrary phase-only filter in the object domain multiplying the input object (the original image), after a Fourier transform we obtain a Fourier domain image and we impose the require Fourier intensity (actually the magnitude), leaving the phase unharmed. An inverse Fourier transform brings us back to the object domain. Since we restrict ourselves to a phase-only filter, we impose the intensity of the input object in this plane. Next, we calculate the Fourier transform and return to the Fourier domain, and so on. This procedure is required since using only the phase of the complex filter that converts the input image exactly to the Fourier image gives poor results. As can be seen, if we impose half of the information (intensity or phase) in both the input and the output domains the procedure converges monotonically.

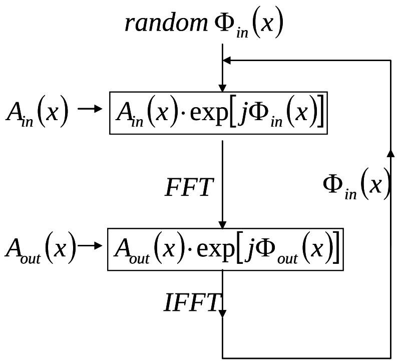

Fig. 1. Block diagram of Gerchberg-Saxton original phase retrieval algorithm.

The Gerchberg-Saxton algorithm is shown in Fig. 1. In later work Gerchberg [9] and Papoulis [10] suggested the use of this method for super-resolution, as shown in Fig. 2. However, both presented relatively simple test cases and assumed the properties of all iterations to be identical (accept when noise reduction was addressed). It is this work of Gerchberg and Papoulis that inspired the authors of this paper.

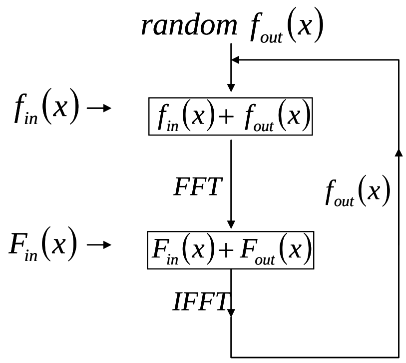

Fig. 2. Block diagram of Gerchberg-Papoulis algorithm. f_{in} and f_{out} are the known and unknown regions in the object plane, respectively, F_{in} and F_{out} are the known and unknown regions in the Fourier image, respectively.

There are of course other methods for obtaining the phase filter, such as Simulated Annealing [11], which ensures that the MSE has indeed a global minimum, but it is time and resources consuming.

C. Super-Resolution by factor of 2

Now we present the dynamic approach used to obtain a factor-2 Super-resolution (i.e. the original image had 2N by 2N pixels but we obtain only an N by N pixels image).

Let us indicate the required image by $g(m,n)$ where $1 \le m,n \le 2N$, and the low-resolution image as $g_{LR}(m,n)$ where $1 \le m,n \le N$. Since the size of the images is not the same we begin by planting zeros between each row and column element of $g_{LR}(m,n)$, thus generating:

$$g_1(m,n) = \begin{cases} g_{LR}\left(\frac{m+1}{2},\frac{n+1}{2}\right) & odd(m\,\&\,n) \\ 0 & otherwise \end{cases} \tag{1}$$

The new image has the same number of pixels as the required high-resolution image.

We assume that a certain portion of the image is known completely, so we may impose it on the new image:

$$g_1(m,n) = g(m,n) \quad m_1 \le m \le m_2, n_1 \le n \le n_2 \tag{2}$$

In the Fourier plane we obtain for the high-resolution image:

$$G(k,l) = \frac{1}{2M \times 2N} \times \sum_{m=1}^{2M}\sum_{n=1}^{2N} g(m,n) exp\left[j2\pi\left(\frac{m \times k}{2M} + \frac{n \times l}{2N}\right)\right] \tag{3}$$

whereas for the low-resolution image, the DFT is

$$G_{LR}(k,l) = \frac{1}{M \times N} \times \sum_{m=1}^{M}\sum_{n=1}^{N} g(m,n) exp\left[j2\pi\left(\frac{m\times k}{M} + \frac{n \times l}{N}\right)\right] \quad (4)$$

Since the low-resolution image contains all the low frequencies of the high-resolution image we may deduce that we can use the following restriction:

$$G_1(k,l) = 4 \times G_{LR}(k,l) \quad k \le M, l \le N \quad (5)$$

Where $G_1(k,l)$ is the DFT of $g_1(m,n)$, and the factor of 4 is required since the size of $g_1(m,n)$ and $g_{LR}(m,n)$ is not the same.

If $m_2 - m_1 + 1 = M$ and $n_2 - n_1 + 1 = N$ then we have only 25% of the data in the object plane and 25% of the data in the Fourier plane. This is less than the required amount stated by the phase retrieval algorithm shown before and therefore the convergence to a minimal MSE is not assured.

In fact, when using the Gerchberg-Papoulis iterative procedure for this case we obtain the following: at first, the MSE decreases monotonically, but after several iterations the requirements are not strict enough to keep the results on the right track and the MSE begins to rise. For this reason we add a new condition to the iterative procedure: halt when MSE reaches local minima.

In addition we noticed that a smaller (better) MSE minimum could be obtained by gradually increasing the frequency domain requirements, i.e., imposing only very low frequencies at the beginning and gradually increasing the frequencies imposed, up to the maximum value given in (5). This is due to the fact that primarily $g_1(m,n)$ is quite different from $g(m,n)$ (yielding a large MSE) and thus imposing a large portion of the Fourier domain at an early stage sets the results way out of track. By doing this we introduce a revised, dynamic version of the Gerchberg-Papoulis algorithm.

The following summarizes the procedure steps:

1. Obtain low-resolution image and set initial low frequencies range to be imposed.
2. Implant zeros between data points to increase image size to $2M \times 2N$.
3. Impose high-resolution portion on image.
4. Perform DFT.
5. Impose low frequencies on DFT.
6. Perform IDFT.
7. Impose high-resolution portion on image.
8. Calculate MSE. If local minimum is obtained then increase the range of low frequencies to impose.
9. If the range of low frequencies has reached the complete range available in the original low-resolution image then stop, else return to step 4.

The block diagram describing the dynamic algorithm is given in Fig. 3 below.

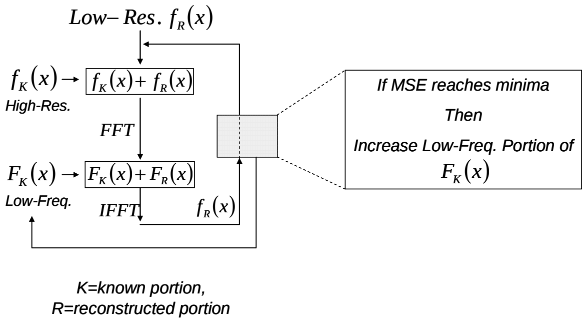

Fig. 3. Block diagram of revised Gerchberg-Papoulis Algorithm.

When using this method we can obtain a sharp image at relatively short processing times. To demonstrate this lets assume that each M by N image requires 1 time unit to process. Thus the original high-resolution image would require 4 time units (because its size is 2M by 2N), however, an image containing only one quarter of the data in high-resolution will require 1.75 time units: 1 time unit for the N by M high resolution and 0.25 time units per each remaining M/2 by N/2 quarters (the zero padding is not relevant in this case). Thus, the processing time is less then half of the original one required if all data were given at high resolution.

III. Simulation Results

The new method was tested on three typical examples. The first, a test rosette containing a variety of frequencies, thus making it easy to view how resolution is improved. The second, a text example showing how barely readable text can be sharpened.

In Fig. 4 we can see the required high-resolution rosette image, whereas Fig. 5 shows the low-resolution rosette. Fig. 6 collaborates the low-resolution data, after padding with zeros to obtain same image size as the high-resolution image, with a portion of the high-resolution image in the first quarter of the image. The first quarter was used since in most cases the higher frequencies lie in the center of the image and we wanted to avoid a biased result.

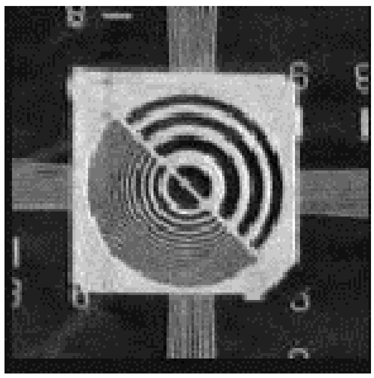

Fig. 4. Original high resolution rosette image containing 128 by 128 pixels.

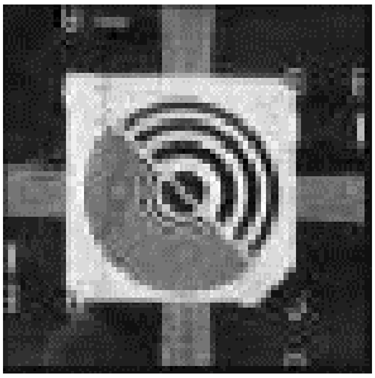

Fig. 5. Original low resolution rosette image containing 64 by 64 pixels.

As can be seen, the high-resolution portion of the image contains both high and low frequencies. The MSE between the first two images is calculated to be 4.337×10^{6} (this value will be normalized to 1).

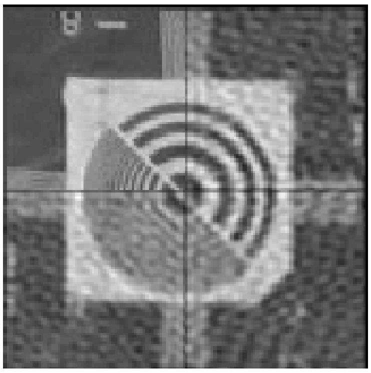

Fig. 6. Low-resolution rosette image after padding with zeros and embedding the high-resolution data in top left corner.

In Fig. 7 one can see the result of the super resolution iterative procedure. MSE is reduced to 1.685×10^{5} (normalized 0.0388) and the correlation coefficient between the high-resolution image and the one obtained is 97.4%.

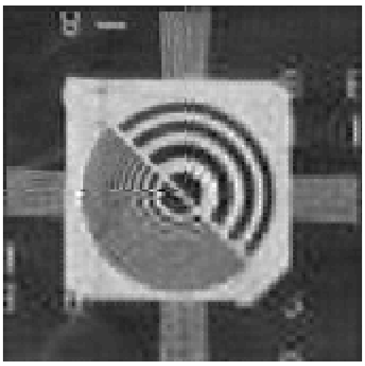

Fig. 7. Reconstructed rosette image

Applying the same algorithms on a text image containing a single word yields the results shown in Fig. 8.

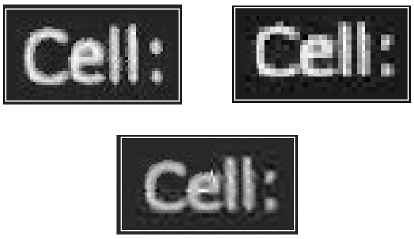

Fig. 8. High resolution word in top-left corner, low resolution word in top-right corner, reconstructed word in bottom line.

As can be seen, in the original high-resolution image (top-left corner) the text is relatively sharp, whereas in the original low-resolution image the text is quite smeared (top-right corner). The original MSE in this case is 3.214×10^{5} and the correlation coefficient is 78.2%. When applying the novel iterative approach one obtains the image shown at the bottom of Fig. 8. The MSE in this case is 8.0088×10^{4} (an improvement by a scale of over 4) and the correlation coefficient is 90.8%, which is much higher than the one obtained by simple reconstruction.

All of the results above assume that the procedure initially uses half of the pixels available in the Fourier domain (e.g., 32) and gradually increases the number of pixels used to the maximum available in the low-resolution image (e.g., 64). It is now time to address the issue of initial number of pixels used, which can be named the initial seed side of the procedure. We use the last test case, reconstructing a text, and sketch both criteria shown above to evaluate the quality of the reconstruction, as a function of the initial seed size.

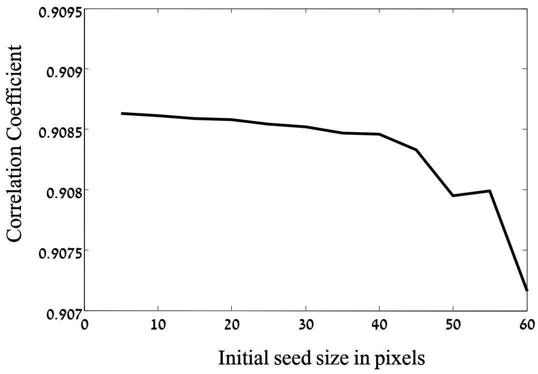

Fig. 9. Correlation coefficient versus initial seed size for a simple text reconstruction example.

As seen in Fig. 9 the final correlation coefficient is almost independent on the initial seed size (less then 0.2% change due to increasing initial seed size from 10 to 60). However, the MSE, shown in Fig. 10, increases as a function of the initial seed size, from 10 pixels to 60 pixels by over 10.1% giving a clear advantage to a smaller seed size.

Finally, the notorious 'Lena' image was put to the test. In Fig. 11 one can see the original low-resolution picture plus zooming on an area not in the top-left quarter (the region in which the high-resolution picture is imposed). In Fig. 12 we can see the reconstructed image. Zooming again on the same area as before it is clear that the original low-resolution pixilation is gone, the tradeoff is a low frequency ripple. This example is important as the MSE improvement by decreasing initial seed size from 60 to 10 pixels is over 43.3%, making the algorithm especially attractive for gray-scale picture reconstruction.

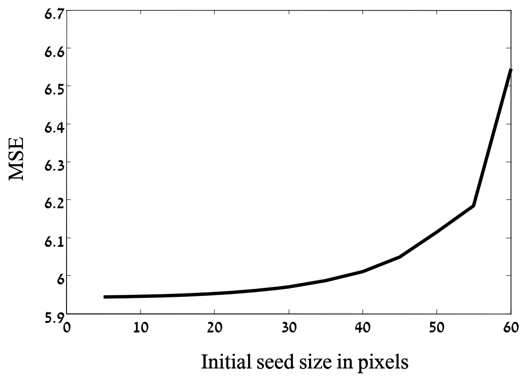

Fig. 10. MSE versus initial seed size for a simple text reconstruction example.

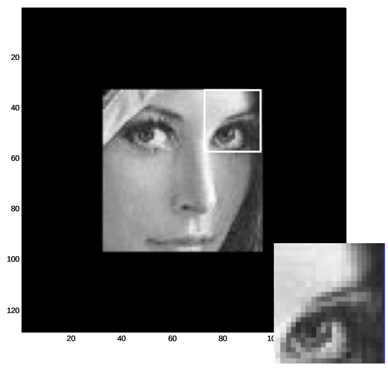

Fig. 11. Low-resolution Lena image before reconstruction.

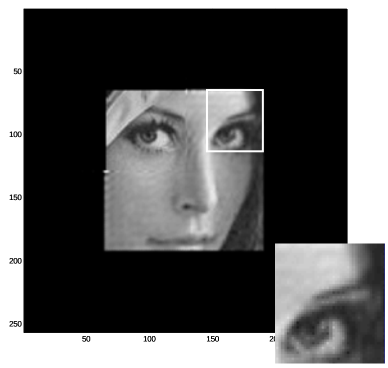

Fig. 12. Reconstructed Lena image.

IV. CONCLUSION

In this paper the authors suggested a novel iterative method for achieving super resolution using a low-resolution image accompanied by a small portion of the high-resolution image, and a dynamic restriction. The new method allows obtaining only a small part of the data with high accuracy and thus saving time while obtaining the images and memory while saving the data before processing. This method may be applicable either as a simple and fast algorithm for slightly improving image content or as a preliminary process before applying advanced digital techniques (e.g., text oriented recognition methods). The suggested procedure is especially useful for reconstructing gray-scale pictures in which only a small portion of the picture is given in high-resolution.

REFERENCES

[1] S. Bellini, "Blind Deconvolution," S. Haykin (Editor), chapter 2, pp. 8-55, 1994.

[2] N.K. Bose, Wavelet-Based Blind Superresolution from Video Sequence and in MRI, PENNSYLVANIA STATE UNIVERSITY – Final Report, 2005.

[3] W. T. Freeman, T. R. Jones, and E. C. PasztorA, Example-Based Super-Resolution, IEEE - Computer Graphics and Applications: Vol. 22, No. 2, pp.56–65, 2002.

[4] H. Chang, D-Y Yeung, and Y. Xiong, Super-Resolution Through Neighbor Embedding, IEEE - Computer Vision and Pattern Recognition (CVPR2004): pp.275-282, 2004.

[5] M. Elad and A. Feuer,. Restoration of a Single Superresolution Image from Several Blurred, Noisy, and Undersampled Measured Images, IEEE Transactions On Image Processing, Vol. 6, No. 12, pp. 1646-1658, 1997.

[6] A. Zomet and S. Peleg, Multi-sensor Super-Resolution, IEEE workshop on Applications of Computer Vision (WACV02), pp. 27-31, 2002.

[7] R.W. Gerchberg and W. O. Saxton, A practical algorithm for determination of phase from image and diffraction plane picture, Optik (Stuttgart), Vol. 35, pp. 237-246, 1972.

[8] J. R. Feinup, Phase retrieval algorithms – a comparison, Applied Optics, Vol. 21, No. 15, pp. 2758-2769, 1982.

[9] R.W. Gerchberg, Super-resolution through error energy reduction, Optica Acta, Vol. 21, No. 9, pp. 709-720, 1974.

[10] A. Papoulis, A new algorithm in spectral analysis and band-limited extrapolation, IEEE transactions on Circuits and Systems, Vol. 22, No. 9, pp. 735-742, 1975.

[11] S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi, Optimization by simulated annealing, Science, Vol. 220, No. 4598, pp. 671-679, 1983.