PHASE RETRIEVAL OF IMAGES FROM ZEROS OF EVEN UNWRAPPED SIGNALS

Styliani Petroudi and Andrew E. Yagle

Dept. of EECS, The University of Michigan, Ann Arbor, MI 48109-2122 aey@eecs.umich.edu

ABSTRACT

The 2-D discrete phase retrieval problem is to reconstruct an image defined at integer coordinates and having known finite spatial extent from the magnitude of its discrete Fourier transform. Most methods for solving this problem are iterative but not POCS, and they tend to stagnate. Recently we developed a new approach that unwrapped the 2-D problem into a 1-D problem with bands of zeros in it, using the Good-Thomas FFT. However, this approach reconstructed the even part of the image much better than the odd part, and it was sensitive to the zero locations. This paper presents a modification of this approach. New features include: (1) an overdetermined problem less sensitive to the zero locations; (2) the solution of a Toeplitz-block-Toeplitz-plus-Hankel-block-Hankel linear system; and (3) details on characteristics of images for which the approach works best.

1. INTRODUCTION

1.1. Basic Problem

The problem of reconstructing an image known to have compact support from its Fourier transform magnitudes arises in several disciplines [1]. The image is reconstructed if the missing Fourier phase is recovered; hence the term ”phase retrieval.” For details of the history and applications of this problem see [1]. Since the image has compact support, its Fourier transform may be sampled in wavenumber. Most images are approximately bandlimited to the extent that they may also be sampled spatially as well. This leads to the discrete version of this problem, in which a discrete-time image known to have finite spatial extent is to be reconstructed from the magnitude of its 2-D discrete Fourier transform (DFT). For details on phase retrieval problems see [2]-[3].

The most common approach for phase retrieval problems is to use an iterative transform algorithm [1], which alternate between the spatial and wavenumber domains. However, these algorithms usually stagnate, failing to converge to a solution. Other approaches require the computationally expensive and extremely unstable numerical operation of tracking zero curves of algebraic functions. We will not attempt to list all approaches here.

1.2. Unwrapping Approach

Recently [4] we proposed a novel approach to this problem. The 2-D phase retrieval problem was slanted and upsampled vertically, and the Good-Thomas FFT or Agarwal-Cooley fast convolution [5] was used to map the 2-D problem into a 1-D problem of reconstructing a 1-D signal consisting of the rows of the image concatenated alternately with bands of zeros. It was noted that the zeros of the z-transform of this 1-D signal are all very close to the unit circle–so close that their angles can be determined from local minima in its (known) 1-D DTFT magnitude. The even and odd parts of the 1-D signal were then reconstructed separately from these zeros, approximated as being on the unit circle.

Although this approach was shown to work in [4], there are several problems: (1) since the even and odd parts are computed directly from the zero locations, they are very sensitive to these zero locations; (2) although the even part of the signal can be reconstructed accurately, direct reconstruction of the odd part turns out to be less accurate; (3) this does not take advantage of the known bands of zeros.

1.3. Contributions of This Paper

This paper presents a modification of our previous approach that addresses the problems noted above, as follows:

1. Instead of computing even and odd parts separately, we now reconstruct the even part first, since its zeros do lie on the unit circle, followed by the odd part;

2. Instead of computing the signal directly from the zero locations, we take advantage of the known bands of zeros in the 1-D signal by formulating this as an overdetermined least-squares interpolation problem, reducing sensitivity to the zero locations;

3. We note that the pseudoinverse of the overdetermined Vandermonde matrix for the interpolation problem with known bands of zeros leads to a Toeplitz block Toeplitz + Hankel block Hankel linear system of equations;

4. Iterative methods such as the Landweber iteration can be used to solve the system quickly using the 2-D FFT. However, we have found that the QR algorithm gives better results, possibly due to conditioning.

2. PROBLEM FORMULATION

2.1. Basics of Phase Retrieval

The 2-D discrete phase retrieval problem is as follows [2]-[3]. $x(i_1, i_2)$ is a discrete-time 2-D image known to be zero except for $|i_1|, |i_2| \leq M/2$. Given knowledge of the 2-D DFT magnitudes $|X(k_1, k_2)|$, compute $x(i_1, i_2)$ or equivalently $X(k_1, k_2)$; hence the term "phase retrieval." The exact order of the 2-D DFT is irrelevant provided each dimension exceeds $2M$, since it may be downsampled or upsampled without difficulty by computing an inverse DFT of the old order, followed by a DFT of the new order. We assume that we are given the autocorrelation $r(i_1, i_2) = x(i_1, i_2) ** x(-i_1, -i_2)$, the inverse DFT of $|X(k_1, k_2)|^2$.

There are two trivial ambiguities in this problem. Clearly if $x(i_1, i_2)$ is a solution then $x(-i_1, -i_2)$ and $-x(i_1, i_2)$ are also solutions. The support condition $|i_1|, |i_2| \leq M/2$ need only be tight enough to prevent any translational ambiguity, so that the image cannot "rattle around" inside this square. Excluding these trivial ambiguities, the 2-D discrete phase retrieval problem almost surely has a unique solution [2]-[3]; we assume this in the sequel.

The 1-D version of this problem has an additional, non-trivial ambiguity. There are almost surely $2^{M/2}$ solutions to an $(M+1)$-point 1-D phase retrieval problem if M is even. This can be seen by noting that the square DFT magnitude is the DFT of the autocorrelation of the 1-D signal. The zeros of the z-transform of the autocorrelation occur in reciprocal and conjugate quadruples (if z is a zero, then $z^*, 1/z, 1/z^*$ are also zeros). Either z and z^*, or their reciprocals, can be assigned to the 1-D signal; since there are $M/2$ quadruples this assignment can be made in $2^{M/2}$ ways (fewer if there are zeros on the unit circle or real axis).

2.2. Reformulation as 1-D Problem

We illustrate the reformulation using the simple example

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} ** \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 11 & 6 \\ 14 & 30 & 14 \\ 6 & 11 & 4 \end{bmatrix}. \tag{1}$$

Note (1) can be written as the 8×3 *cyclic* convolution

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ** \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 30 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 6 \\ 11 & 0 & 0 \\ 0 & 4 & 4 \\ 11 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 14 \end{bmatrix} \tag{2}$$

Note that both the image pixel values and the autocorrelation lags have been "slanted," but this does not alter the operation of autocorrelation. Also note the upsampling (inserting bands of zeros vertically between image pixel values); this also does not alter the operation of autocorrelation. The general problem for an $M \times M$ image can be reformulated as a $(N^2-1) \times N$ 2-D cyclic convolution, where $N \geq 2M-1$, which ensures that there is no aliasing. The reason for using (N^2-1) is explained next.

2.3. Unwrapping 2-D to 1-D using Agarwal-Cooley

The Agarwal-Cooley fast convolution algorithm maps an $N_1 \times N_2$-order 2-D cyclic convolution to an $N_1 N_2$-point 1-D cyclic convolution using a residue number system (RNS) mapping $x(i_1, i_2) \to x(i)$ where

$$i = i_1 (mod\, N_1); \quad i = i_2 (mod\, N_2), \tag{3}$$

where N_1 and N_2 are relatively prime. Since N^2-1 and N are relatively prime, we can use Agarwal-Cooley to rewrite the 2-D cyclic convolution (2) as a 24-point 1-D cyclic convolution $x(n) * x(-n) = r(n)$, viz.

$$\{1,2,0,3,4,0\ldots0\} * \{0\ldots0,4,3,0,2,1\}$$
$$= \{30,14,6,11,4,0\ldots0,4,11,6,14\} \tag{4}$$

The zero-padding in (4) makes this cyclic convolution equivalent to a linear convolution. Note the bands of single zeros in the unwrapped image, which is formed by concatenating rows of the original image, alternating with bands of zeros. Since the 1-D autocorrelation is also an unwrapping of the 2-D autocorrelation, the magnitude of the DTFT of the unwrapped image is known.

Alternately, we may use the Good-Thomas FFT maps an $N_1 \times N_2$-orderhe FFT to an $N_1 N_2$-point 1-D DFT using an RNS mapping; see [4] for details.

3. SOLUTION OF UNWRAPPED PROBLEM

3.1. Problem with Previous Approach

In [4] we proposed the reconstruction of the unwrapped signal $x(n)$ from its known autocorrelation $r(n)$ by observing that the zeros of the z-transform $R(z)$ of the latter are very close to the unit circle. At the frequencies ω_i corresponding to the angles of these zeros, the known DTFT magnitude $|X(e^{j\omega})|$ is near zero, so that the real and imaginary parts must both also be near zero. Recalling that the even $x_e(n)$ and odd $x_o(n)$ parts of $x(n)$ are defined as

$$x_e(n) = (x(n)+x(-n))/2 = dtft^{-1}[Re[X(e^{j\omega})]]; \tag{5a}$$

$$x_o(n) = (x(n)-x(-n))/2 = dtft^{-1}[Im[X(e^{j\omega})]]/j, \tag{5b}$$

these can be reconstructed from these frequencies using

$$X_e(z) = \prod (z - e^{-j\omega_i}) \tag{6}$$

and similarly for $X_o(z)$. Some complications in reconstructing $X_o(z)$ and determining scale factors were noted in [4].

All of the material to follow is entirely new.

We have observed that this approach reconstructs $x_e(n)$ well, but reconstructs $x_o(n)$ not nearly as well as $x_e(n)$. The reason is shown in Fig. 1. Fig. 1a plots the zeros of

$$x(n) = \{2, \underbrace{0\ldots0}_{9\,zeros}, 1, 4, \underbrace{0\ldots0}_{9\,zeros}, 3\} \tag{7}$$

while Fig. 1b plots the zeros of its even part and Fig. 1c plots the zeros of its odd part. Note that the zeros of $x(n)$ are close to the unit circle, but the zeros of $x_e(n)$ lie *on* the unit circle, with angles very close to the angles of the zeros of $x(n)$. However, the zeros of $x_o(n)$ are farther away from the unit circle than those of $x(n)$ or $x_o(n)$.

The reason for this is illustrated in Fig. 2, which show a larger unwrapped image. Fig. 2a shows $x(n)$ (note the zero bands), Fig. 2b shows $x_e(n)$ (note the similarity to $x(n)$, even though it is only the even part), and Fig. 2c shows $x_o(n)$. Due to their similar appearance, it is not surprising that the zeros of both $x(n)$ and $x_e(n)$ lie close to the unit circle, while the zeros of $x_o(n)$ may not. We have observed this repeatedly for images with non-negative pixels (which is the case of practical interest) and especially for images whose pixel values lie in a bounded range, for which this approach works best.

3.2. Summary of New Approach

Although $x_e(n)$ can be reconstructed using (6), this does not take advantage of the bands of zeros in $x_e(n)$ which are apparent in Fig. 2b. This motivates the following approach:

1. Threshold the known DTFT square magnitude $|X(e^{j\omega})|^2$ to determine the frequencies ω_i where $|X(e^{j\omega_i})| \approx 0$. As Fig. 1 shows, $e^{j\omega_i}$ (*on* the unit circle) are good approximations to the actual zeros of $x_e(n)$;

2. Reconstruct $x_e(n)$ by solving the *overdetermined*

$$\begin{bmatrix} 1 & 1 \\ \cos(\omega_1) & \cos(\frac{M^2}{2}\omega_1) \\ \vdots & \vdots \\ \cos(\omega_L) & \cos(\frac{M^2}{2}\omega_L) \end{bmatrix} \begin{bmatrix} x_e(1) \\ \vdots \\ x_e(\frac{M^2}{2}) \end{bmatrix} = \begin{bmatrix} |X(e^{j0})| \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{8}$$

   where L is the number of frequencies used and $M^2/2$ is the number of unknown *nonzero* values of $x_e(n)$;

3. Using the computed $x_e(n)$ in (5a), compute

$$Im[X(e^{j\omega})]^2 = |X(e^{j\omega})|^2 - Re[X(e^{j\omega})]^2 \tag{9}$$

   and determine the frequencies $\tilde{\omega}_i$ of its zero crossings;

4. Reconstruct $x_o(n)$ by solving an overdetermined system analogous to (8) but with sines rather than cosines;

5. Determine the scale factor for $x_o(n)$ using a fit to the known values of $Im[X(e^{j\omega})]^2$.

This approach offers the following advantages over [4]:

1. It takes advantage of the known bands of zeros in $x_e(n)$ to produce an overdetermined problem which is less sensitive to errors in the angles of the zeros;

2. It requires only that the zeros of $x_e(n)$ lie close to the unit circle (which they do), not the zeros of $x_o(n)$ to lie close to the unit circle (which they may not);

3. It avoids the scale factor computation in [4].

A disadvantage is that the reconstruction of $x_o(n)$ is affected by errors in the reconstruction of $x_e(n)$. However, the latter can be reconstructed much more accurately, so the independence of the reconstructions in [4] is not helpful.

3.3. Computation of Solution

The least-squares solution to the overdetermined problem (8) uses the pseudoinverse of the matrix V in (8); it solves

$$(V^T V)\hat{x}_e = V^T[|X(e^{j0})|, 0\ldots0]' = |X(e^{j0})|[1, 1\ldots1]'. \tag{10}$$

The $(i,j)^{th}$ element of $V^T V$ has the form

$$2(V^T V)_{i,j} = 2\sum_{n=1}^{L}\cos(i\omega_n)\cos(j\omega_n)$$

$$= \sum_{n=1}^{L}\cos((i-j)\omega_n) + \sum_{n=1}^{L}\cos(i+j)\omega_n) \tag{11}$$

except that rows and columns corresponding to the bands of zeros in $x_e(n)$ have been deleted. The resulting matrix is Toeplitz-block-Toeplitz plus Hankel-block-Hankel.

Hence (8) can be solved iteratively using the Landweber iteration, since the matrix-vector products required at each iteration can be computed quickly using the 2-D FFT. Or use variations on the multichannel Levinson algorithm.

A analogous identity can be used to write the system to be solved for $x_o(n)$ as Toeplitz-block-Toeplitz minus Hankel-block-Hankel. Similar comments apply.

Another approach that was tried was to omit the first row of (8). Then the problem is to compute the eigenvector associated with the minimum eigenvalue of $V^T V$, i.e., the minimum singular vector of V (omitting its first row). This can be computed by using the power method to compute the maximum singular value σ_{max}, and then using the power method again to compute the maximum singular vector of $\sigma^2_{max} I - V^T V$, which is the desired vector. However, this did not give results that were as good.

3.4. Summary of Results

Due to lack of space, we summarize our results here:

1. This approach works well when the image pixel values do not vary too much (on a percentage basis), as in Fig. 2. For such images, the unwrapped 1-D signal tends to have zeros closer to the unit circle than images with widely varying pixel values. Since many images do in fact have a narrow range of pixel values, this is not a significant problem;

2. We obtain better results using the QR decomposition of the matrix in (8) than by using the pseudoinverse solution. We attribute this to numerical conditioning problems. However, this increases the computation significantly;

3. Since the zeros of the autocorrelation $R(z)$ lie very close to the unit circle, their approximate locations are indicated by dips in the magnitude of the DTFT of $R(z)$. However, we may use these approximate locations as an initialization of Newton's method to find the exact zeros of $R(z)$, and their angles. Use of the angles of the actual zero locations improves performance, again at the price of increased computation;

4. Due to lack of space, we do not present detailed reconstructions here. These will be presented at ICASSP and in the full-length paper.

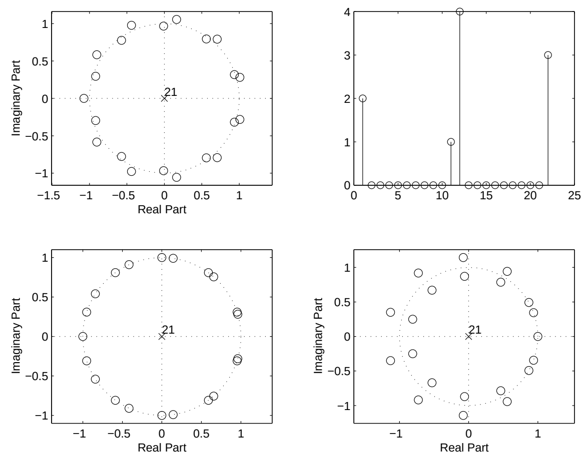

Fig. 1. Figs. 1a,1b,1c: Zeros of $x(n), x_e(n), x_o(n)$

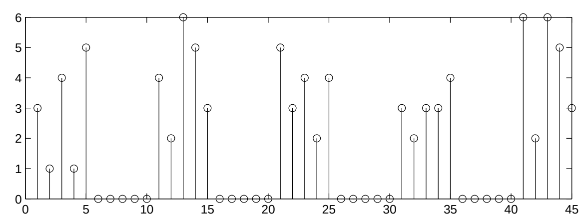

Fig. 2. Fig. 2a: Unwrapped signal

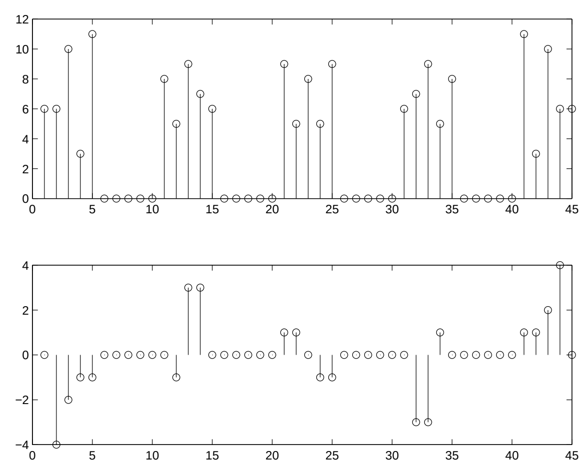

Fig. 3. Figs. 2b,2c: Even and odd parts

4. REFERENCES

[1] J.C. Dainty and J.R. Fienup, "Phase retrieval and image reconstruction for astronomy," in *Image Recovery: Theory and Application*, ed. by H. Stark, Academic Press, San Diego, 1987, pp. 231-275.

[2] M.H. Hayes, "The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform," IEEE Trans. ASSP 30(2), 140-154 (1982).

[3] M.H. Hayes, J.S. Lim, and A.V. Oppenheim, "Signal reconstruction from phase or magnitude," IEEE Trans. ASSP 28(6), 672-680 (1980).

[4] S. Petroudi and A.E. Yagle, "Phase retrieval of images from the zeros of 1-D signals obtained using the Good-Thomas FFT," Proc. ICASSP, June 2000, Istanbul.

[5] R.E. Blahut, *Fast Algorithms for Digital Signal Processing*, Addison-Wesley, NY, 1985.