HOW TO TAKE ADVANTAGE OF ALIASING IN BANDLIMITED SIGNALS

Patrick Vandewalle[1] *, Luciano Sbaiz[1], Joos Vandewalle[2] † and Martin Vetterli[1,3]

[1]LCAV - School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL) - CH-1015 Lausanne, Switzerland
[2]SCD - Department of Electrical Engineering (ESAT)
Katholieke Universiteit Leuven - Kasteelpark Arenberg 10, B-3001 Heverlee, Belgium
[3]Department of Electrical Engineering and Computer Science
UC Berkeley - Berkeley CA94720

ABSTRACT

In signal processing systems, aliasing is normally treated as a disturbing signal. That motivates the need for effective analog, optical and digital anti-aliasing filters. However, aliasing conveys also valuable information on the signal above the Nyquist frequency. Hence, an effective processing of the samples, based on a model of the input signal, would allow to virtually increase the sampling frequency using slower and cheaper converters. In this paper, we present such an algorithm for bandlimited signals that are sampled below twice the maximum signal frequency. Using a subspace method in the frequency domain, we show that these signals can be reconstructed from multiple sets of samples. The offset between the sets is unknown and can have arbitrary values. This approach can be applied to the creation of super-resolution images from sets of low resolution images. In this application, registration parameters have to be computed from aliased images. We show that parameters and high resolution images can be computed precisely, even when high levels of aliasing are present on the low resolution images.

1. INTRODUCTION

In general, we say that a signal can be perfectly reconstructed from its samples if it is bandlimited and the sampling frequency satisfies the Nyquist criterion, i.e. it is larger than twice the maximum signal frequency. If the signal is not bandlimited or the sampling frequency is too low, the sampled signal is aliased, and perfect reconstruction is not possible.

Vetterli et al. [1] showed that perfect reconstruction is also possible for signals with finite rate of innovation. Certain types of non-bandlimited signals (like streams of Diracs, piecewise polynomials, etc.) can be reconstructed from a finite set of samples. Vaidyanathan [2] considers some other sampling schemes for non-bandlimited signals, like the use of different sampling kernels. An overview of the current state in sampling is given by Unser [3].

In this paper, we derive some results for bandlimited signals that are sampled below twice the maximum frequency of the signal. We will use multiple sets of regular samples to reconstruct the original signal exactly. A similar problem for discrete-time signals is solved by Marziliano et al. [4] using combinatorial optimization methods.

2. PROBLEM STATEMENT

We define $x(t)$ as a periodic, bandlimited and continuous-time signal with period 1 and maximum frequency L (since we consider a Fourier series L is integer). Its continuous time Fourier series is called $X[l]$ (with $-L \leq l \leq L$).

$x(t)$ is regularly sampled at frequency K, with K integer-valued. This results in a discrete-time signal $y_0[n] = x[n/k]$. Its discrete-time Fourier transform $Y_0[k]$ can be represented as a function of the continuous-time Fourier coefficients $X[k]$ as

$$Y_0[k] = \sum_{i=-\infty}^{\infty} X[k+iK] \text{ with } -\left\lfloor \frac{K}{2} \right\rfloor \leq k \leq \left\lfloor \frac{K-1}{2} \right\rfloor. \tag{1}$$

$Y_0[k]$ is a periodic signal with period K, so it is sufficient if we only look at one period of this signal.

If we want to reconstruct $x(t)$ from its samples, we need to know its Fourier coefficients $X[k]$. If the sampling frequency satisfies the Nyquist criterion ($K > 2L$), $X[k+iK]$

is only different from zero for $i = 0$ and

$$Y_0[k] = X[k] \text{ with } -\left\lfloor \frac{K}{2} \right\rfloor \leq k \leq \left\lfloor \frac{K-1}{2} \right\rfloor. \quad (2)$$

However, if the Nyquist criterion is not satisfied ($K \leq 2L$), the copies of the continuous spectrum overlap in the sampled spectrum. In equation (1) $X[k + iK]$ is nonzero for multiple values of i and therefore the Fourier coefficients $X[k]$ of the continuous signal cannot be derived immediately.

Now we take a second (regular) set of samples $y_1[n]$ at the same sampling rate K, with an unknown offset t_1 from the first set $y_0[n]$: $y_1[n] = x\left(\frac{n}{K} + t_1\right)$ with $t_1 \in [0, 1)$. Again, we can write its Fourier coefficients $Y_1[k]$ as a function of $X[k]$:

$$Y_1[k] = \sum_{i=-\infty}^{\infty} W_{t_1}^{k+iK} X[k + iK]$$

$$= W_{t_1}^{k} \sum_{i=-\infty}^{\infty} W_{t_1}^{iK} X[k + iK], \quad (3)$$

with $W_\alpha = e^{j2\pi\alpha}$ and $-\lfloor K/2 \rfloor \leq k \leq \lfloor (K-1)/2 \rfloor$. This gives us K new equations in $X[k]$, but it adds also a new unknown t_1.

We can reformulate equation (1) and (3) using vector notation and using the period from 0 to $K - 1$ instead of $-K/2$ to $K/2$:

$$\begin{bmatrix} Y_m[0] \\ Y_m[1] \\ \vdots \\ Y_m[K-1] \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & W_{t_m} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & W_{t_m}^{K-1} \end{bmatrix}$$

$$\sum_{i=-\infty}^{\infty} W_{t_m}^{iK} \begin{bmatrix} X[iK] \\ X[iK+1] \\ \vdots \\ X[(i+1)K-1] \end{bmatrix}$$

$$\Updownarrow$$

$$\mathbf{Y}_m = \mathbf{D}_m \sum_{i=-\infty}^{\infty} W_{t_m}^{iK} \mathbf{X}_{iK}^{(i+1)K-1}$$

$$\Updownarrow$$

$$\mathbf{D}_m^{-1} \mathbf{Y}_m = \sum_{i=-\infty}^{\infty} W_{t_m}^{iK} \mathbf{X}_{iK}^{(i+1)K-1}, \quad (4)$$

with t_m the relative offset from the set $y_0[n]$ ($t_0 = 0$) and m the index of the sample set ($0 \leq m < M$). Because the original signal $x(t)$ is bandlimited, $\mathbf{X}_{iK}^{(i+1)K-1} \neq 0$ only for a finite number S of values i:

$$-\left\lfloor \frac{L+K-2}{K} \right\rfloor \leq i \leq \left\lfloor \frac{L-1}{K} \right\rfloor \quad (5)$$

with $\lfloor x \rfloor$ the largest integer value smaller than x. This allows us to compute S as

$$S = \left\lfloor \frac{L+K-2}{K} \right\rfloor + \left\lfloor \frac{L-1}{K} \right\rfloor + 1. \quad (6)$$

We can now make the important observation that the modified sets of samples $\mathbf{D}_m^{-1}\mathbf{Y}_m$ *are all linear combinations of a set of S vectors* $\mathbf{X}_{iK}^{(i+1)K-1}$. Intuitively, it can already be seen that by taking the number of sampling sets M large enough, we will have enough equations to compute the Fourier coefficients $X[k]$ as well as the unknown offsets t_m.

3. SOLUTION USING SUBSPACES

3.1. Offset estimation

As mentioned before, the modified sets of Fourier coefficients of the samples are linear combinations of S vectors $\mathbf{X}_{iK}^{(i+1)K-1}$. Therefore, they belong to an S-dimensional subspace of the K-dimensional space and the subspace matrix

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_0 & \mathbf{D}_1^{-1}\mathbf{Y}_1 & \cdots & \mathbf{D}_{M-1}^{-1}\mathbf{Y}_{M-1} \end{bmatrix}$$

$$= \begin{bmatrix} Y_0[0] & Y_1[0] & \cdots & Y_{M-1}[0] \\ Y_0[1] & W_{t_1}Y_1[1] & \cdots & W_{t_{M-1}}Y_{M-1}[1] \\ Y_0[2] & W_{t_1}^2 Y_1[2] & \cdots & W_{t_{M-1}}^2 Y_{M-1}[2] \\ \vdots & \vdots & & \vdots \\ Y_0[K-1] & W_{t_1}^{K-1}Y_1[K-1] & \cdots & W_{t_{M-1}}^{K-1}Y_{M-1}[K-1] \end{bmatrix}$$

$$(7)$$

has rank S (if $M \geq S$). We assume here that the rank is not lower than S, meaning that there are no two sets of samples with offsets t_i, t_j for which $(t_i - t_j)K \in \mathbb{Z}$. This would be a degenerate case, because two such sets contain the same samples.

If the number of sampling sets M is at least $S + 1$, it is possible to find the relative offsets as the set of parameter values for which the $(S + 1)$-th singular value of $\mathbf{Y}$ equals zero.

3.2. Computation of the Fourier coefficients

Once the relative offsets t_m are known, it is very easy to compute the Fourier coefficients $X[k]$ of the original signal. Each of the equations in (3) is an equation in at most S unknown Fourier coefficients. Because we have $M \geq S+1$ sets of samples, there are also $S + 1$ such equations in the same unknowns. Therefore, the unknown Fourier coefficients are the solution of a linear set of equations.

4. INTERPRETATION

The unknown matrices $\mathbf{D}_m^{-1}$ by which the vectors $\mathbf{Y}_m$ are multiplied, are rotation matrices. They do not modify the

energy in the different vectors (or equivalently, the Frobenius norm of the matrix $\mathbf{Y}$), but they rather align the different vectors such that they are contained in the S-dimensional subspace.

As described in Section 3, we are looking for the t_1, t_2,..., t_{M-1} for which the $(S+1)$-th singular value is 0. Next to the global minimum, this function has also many local minima. It is not possible to apply a standard minimization algorithm (like gradient descent) to the problem, because it will get stuck in one of the local minima. A possibility would be to compute the $(S+1)$-th singular value on a regular grid of K^S values and apply a nonlinear minimization algorithm to the minimum of those values. Obviously, for large values of S, this is unfeasible.
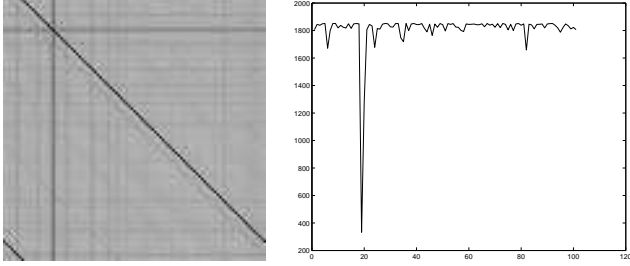
(a) Variation over two offsets.   (b) Variation over one offset.

Fig. 1. Third singular value for different offset values, with $S = 2$ (different values for the two offsets (a), and different values for the second offset if the first is known(b)). This function has many local minima, which makes it difficult to find the global minimum. The gray level indicates the value (white stands for high values, black for zero). The values have been scaled down for display purposes.

It is therefore very important to understand the structure that is present in this multivariate function. As we look at the $(S + 1)$-th singular value as a function of the different t_m (see Figure 1 for an example with $S = 2$), we can also see distinct horizontal, vertical and diagonal lines that go through the global minimum at the intersection of the horizontal, vertical and diagonal lines. These lines correspond to an alignment of $\mathbf{Y}_0$ and $\mathbf{Y}_1$, $\mathbf{Y}_0$ and $\mathbf{Y}_2$, and $\mathbf{Y}_1$ and $\mathbf{Y}_2$, respectively. As these lines can be observed in the plots, they indicate that it is often possible to find most of the pairwise registrations independently, using arbitrary values for the other t_m's.

5. COMPUTATIONAL ISSUES

5.1. Offset estimation

The minimization of the $(S + 1)$-th singular value of $\mathbf{Y}$ can also be seen as a search for the values of t_m ($m =$

$1, 2, ..., S$) for which

$\mathbf{A} = \mathbf{Y}^*\mathbf{Y}$

$$
= \begin{bmatrix}
\mathbf{Y}_0^*\mathbf{Y}_0 & \mathbf{Y}_0^*\mathbf{D}_1^{-1}\mathbf{Y}_1 & \cdots \mathbf{Y}_0^*\mathbf{D}_M^{-1}\mathbf{Y}_M \\
\mathbf{Y}_1^*\mathbf{D}_1^{*-1}\mathbf{Y}_0 & \mathbf{Y}_1^*\mathbf{Y}_1 & \cdots \mathbf{Y}_1^*\mathbf{D}_1^{*-1}\mathbf{D}_M^{-1}\mathbf{Y}_M \\
\vdots & \vdots & \ddots \vdots \\
\mathbf{Y}_M^*\mathbf{D}_M^{*-1}\mathbf{Y}_0 & \mathbf{Y}_M^*\mathbf{D}_M^{*-1}\mathbf{D}_1^{-1}\mathbf{Y}_1 & \cdots \mathbf{Y}_M^*\mathbf{Y}_M
\end{bmatrix}
$$
(8)

is rank deficient. Instead of computing singular values, we can then also simply compute the determinant of $\mathbf{A}$. This will require much less computations, but it will also be numerically less stable, as we do not use all of the available information.

A more stable solution is to compute the smallest eigenvalue of $\mathbf{A}$, and thus its smallest singular value using the inverse power method as described by Strang [5]. Due to the fact that the $(S + 1)$-th singular value is much smaller than the S-th singular value, this method converges very fast. It requires only a few iterations in which a linear system with matrix $\mathbf{A}$ has to be solved.

5.2. Computation of the Fourier coefficients

The set of equations from which the Fourier coefficients are computed, is generally overdetermined. Therefore, it is better to use the least squares solution of the equations. This is the optimal solution when the measurements are perturbed by additive Gaussian noise.

We assume that the offset-related values ($t_m \bmod 1/K$) are uniformly spread over the interval $[0, 1/K)$. This means that not all sample sets $\mathbf{Y}_m$ are at almost the same position, because this will make the reconstruction problem very badly conditioned. The samples would be periodically very close to each other, leaving large gaps with no information in between.

6. APPLICATION: SUPER-RESOLUTION

Super-resolution imaging is one of the typical applications we think of. The goal of super-resolution is to reconstruct a high resolution image from a set of low resolution images of the same scene [6]. If the low resolution images are aliased (which would typically be the case if no special processing is performed), we can apply our algorithm to it.

The idea described for one-dimensional signals in the previous sections can be easily extended to two (or more) dimensions. The offset is now a combination of a horizontal and a vertical shift between the images. Of course, the computational issues mentioned in Section 5 become even more important here, as the minimum number of images for a (minimally) aliased image is five, instead of three for one-dimensional signals. A result with this algorithm is shown in Figure 2.

Fig. 2. Results of the super-resolution algorithm. The (255x255) original image is perfectly reconstructed (right) from a set of five low resolution (128x128) images (left).

7. DISCUSSION AND FUTURE WORK

The algorithm described above can be applied to any type of aliased bandlimited signals, for which different sets of samples are taken. The undersampling factor can be estimated from the evolution of the different singular values of the modified measurement matrix. We can build a matrix from all the sampled sets, compute the singular value decomposition for different offsets and check for which S $\sigma_S \gg \sigma_{S+1}$. This determines the minimum number of sampling sets required for offset estimation and signal reconstruction.

As previously stated, this idea relies totally on the presence of the signal energy beyond the Nyquist frequency in the measured signal. If this signal energy (aliasing) has been removed from the signal prior to the measurement, the information is irremediably lost and the algorithm will not bring any improvements.

Most real signals are not periodic, but have finite extent (or they are sampled over a finite range). Our method can still be applied to such signals, by using the periodic extension of the signals. In that case, there are non-overlapping parts at the borders. It would be very interesting to study the performance of our method for different amounts of overlap between the different signals.

Our method performs well under noisy conditions. This could be expected, as the singular value decomposition is an algorithm that is not very sensitive to noise. Using a least squares method, the same is true for the reconstruction. This noise insensitivity is important in practice, as (almost) all signal measurements will be corrupted by noise.

Throughout this paper, we assumed sampling using a Dirac sampling kernel. No impulse response or point spread function was taken into account. This is a good approximation for real systems, as long as the signals are sampled at (relatively) precise locations. As soon as the impulse response becomes important (e.g. averaging over a certain time/area) we also need to take it into account. This can be done by multiplying $X[k]$ by $H[k]$ in all the equations. The effect of the sampling kernel can then be cancelled at the end by dividing the result by $H[k]$. Of course, if the sampling kernel is too smooth, the signal is low-pass filtered, and all aliasing is removed from the measured signal.

More work needs to be done also on the computational efficiency of the algorithm. On the one hand, the algorithm will rarely find the global minimum if it is immediately started from an arbitrary position, without prior evaluation at a certain number of values. On the other hand, using an exhaustive search, the algorithm would be too computationally intensive (with the complexity even going to infinity if we want to search the continuous shifts exhaustively).

8. CONCLUSIONS

We presented a new frequency domain approach to reconstruct signals from multiple sets of aliased samples. The sets of samples are regular, but the offset between the different sets is unknown. This offset is first computed using a subspace approach. Afterwards, the Fourier coefficients of the original signal can be solved from a set of linear equations. Some computational issues were handled to avoid solutions to be trapped in local minima and to improve computational efficiency. The ideas are then applied to super-resolution imaging. The results show the validity of the method.

9. REFERENCES

[1] Martin Vetterli, Pina Marziliano, and Thierry Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, June 2002.

[2] P. P. Vaidyanathan, "Sampling theorems for non bandlimited signals: theoretical impact and practical applications," in *Proceedings 4th International Conference on Sampling Theory and Applications*, May 2001.

[3] Michael Unser, "Sampling—50 Years after Shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, April 2000.

[4] Pina Marziliano and Martin Vetterli, "Reconstruction of irregularly sampled discrete-time bandlimited signals with unknown sampling locations," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3462–3471, December 2000.

[5] Gilbert Strang, *Linear Algebra and Its Applications*, Saunders College Publishing, third edition, 1988.

[6] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, May 2003.