Super-resolution of interleaved sampled signals

D. H. Justice, H. J. Trussell
Dept. of Electrical
and Computer Engineering
North Carolina State University
Raleigh, NC 27695-7911

R. E. Hartwig
Dept. of Mathematics
North Carolina State University
Raleigh, NC 27695-8205

Abstract

A single ADC is often used to digitize several channels. In this application, the channels are sampled sequentially in time. This results in signals that are correlated but are not sampled at the same time. Bandlimited interpolation can be used to obtain signals with the same sampling times. However, methods that were recently developed for digital cameras can be used to achieve resolution beyond the Nyquist limit of the sampling rate of a single channel. This work demonstrates the effectiveness of this method on stereo signals and shows it to be robust to estimations of the modelling parameters.

1 Introduction

There are several sampling devices that use a single analog-to-digital converter when sampling multiple channels. The methodology for this is to switch sequentially through the channels, digitizing each one in turn. The sampling of the channels, then, does not occur simultaneously. Operations that require comparisons between channels, e.g., correlations, should take this timing difference into account. The common way to do this is to use bandlimited interpolation on the channels and resample the signals on the same time base.

If the channels are correlated and not bandlimited, it is possible to obtain resolution beyond the bandlimit. This is a common problem in color image processing, where the three colors represent three correlated channels. In addition, video imaging provides another application where images at different times can be considered as different channels. Some previous work in this area is discussed in [1, 2]. A monochrome image that is sampled by multiple CCD arrays can be considered as a multichannel signal with interleaved samples [3]. This work uses a least squares method, derived from an image processing technique to obtain high resolution color images from a digital camera [4], to estimate interpolated samples. This is the only known demosaicking methods that can be applied to this problem, as the other published methods rely on unique properties of color images.

2 Mathematical Statement of Problem

For ease of notation, we will formulate the problem using a discrete system. We assume that the full-resolution data recording system for a single channel is modeled by

$$\mathbf{y}_{N\times 1} = \mathbf{H}_{N\times N}\ \mathbf{x}_{N\times 1} + \eta_{N\times 1}, \tag{1}$$

where $\mathbf{x}$ is the original signal that is bandlimited to the sampling rate that produced the samples in the vector, $\mathbf{H}$ is the matrix that represents the distortion of the channel, an impulse response of length $K << N$, and η represents signal independent noise. For ease of computation, we pad the signal and use circular convolution. This lets us use the DFT and makes the matrices square.

For a P-channel system, each channel can be modeled as eq.(1). If the channels are independent, that is, there is no cross-talk, the multi-channel system can be represented using stacked notation.

$$\mathbf{y}_{NP\times 1} = \begin{bmatrix} y_0(0) \\ y_0(1) \\ \vdots \\ y_0(N-1) \\ y_1(0) \\ y_1(1) \\ \vdots \\ y_1(N-1) \\ y_2(0) \\ \vdots \\ y_{P-1}(N-1) \end{bmatrix} = \mathbf{H}_F \begin{bmatrix} x_0(0) \\ x_0(1) \\ \vdots \\ x_0(N-1) \\ x_1(0) \\ x_1(1) \\ \vdots \\ x_1(N-1) \\ x_2(0) \\ \vdots \\ x_{P-1}(N-1) \end{bmatrix}, \tag{2}$$

The matrix $\mathbf{H}_{NP \times NP}$ has the form

$$\begin{bmatrix} \mathbf{H}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{H}_P \end{bmatrix},$$

where $\mathbf{H}_k$ is $N \times N$. If the blurs can be approximated by the same matrix, $\mathbf{H}_k = \mathbf{H}$ for all k then the large matrix can be represented by the Kronecker (outer) product

$$\mathbf{H}_F = \mathbf{I}_P \otimes \mathbf{H},$$

where subscript on the $\mathbf{I}_P$ represents the size of the identity matrix.

The sequential sampling with equally spaced samples can be modeled by a matrix multiply. The 1:2 sampling matrix for a one-dimensional signal is represented by diagonal matrices

$$\mathbf{C}_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

or

$$\mathbf{C}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

The diagonal matrices can be represented more concisely by using

$$\mathbf{C}_0 = diag(1, 0, 1, 0, \cdots, 1, 0)$$

and

$$\mathbf{C}_1 = diag(0, 1, 0, 1, \cdots, 0, 1)$$

The extension to 1:M sampling, where N is an integral multiple of M, is straightforward with every M^{th} diagonal element equal to unity. Note that we have the constraint

$$\sum_{k=0}^{M-1} \mathbf{C}_k = \mathbf{I} \tag{3}$$

When we combine this with the stacked notation, the full sampling matrix is given by

$$\mathbf{C}_F = diag(\mathbf{C}_0, \mathbf{C}_1, \dots, \mathbf{C}_{P-1}), \tag{4}$$

The multichannel model can now be written

$$\mathbf{y} = \mathbf{C}_F[\mathbf{H}_F \mathbf{x}_F + \eta_F].$$

where the subscript F indicates the stacked notation.

Finally, we can state the problem as

$$minimize \quad \Phi(\mathbf{D}) = E\{||\mathbf{x} - \mathbf{D}\mathbf{y}||^2\}$$

with respect to the deconvolution matrix $\mathbf{D}$. To solve this we will use the following matrices.

$$\mathbf{H}_F = \mathbf{I} \otimes \mathbf{H}$$

We assume that all channels responses are equal.

$$\mathbf{W}_F = \mathbf{I} \otimes \mathbf{W}$$

This is the DFT matrix. It takes the DFT of each channel.

$$\mathbf{K}_{xx\lambda} = \mathbf{K}_\lambda \otimes \mathbf{K}_{xx}$$

This is the cross-correlation matrix of the channels. It is assumed separable for computational purposes.

$$\mathbf{K}_{\eta\eta} = \mathbf{I} \otimes \sigma_\eta^2 \mathbf{I}.$$

The noise is assumed signal-independent within the channels and independent across channels.

3 Solution

The solution to the problem is given by

$$\mathbf{D} = \mathbf{K}_{xx\lambda} \mathbf{H}_F^T \mathbf{C}^T [\mathbf{C}\mathbf{H}_F \mathbf{K}_{xx\lambda} \mathbf{H}_F^T \mathbf{C}^T + \mathbf{C}\mathbf{K}_{\eta\eta}\mathbf{C}^T]^\dagger \tag{5}$$

where the $\dagger$ indicates the pseudoinverse. In order to compute this numerically, it is advantageous to transform the problem to the Fourier domain. Because we have assumed circular convolutions, the DFT matrix diagonalizes the system response matrix and the correlation matrices. The transform of the subsampling matrices is a block structure with a phase factor. Assume the first sampling starts at the origin; then the Fourier transform of the k^{th} shifted sampling in one dimension is

$$\mathbf{W}\mathbf{C}_k\mathbf{W}^{-1} = \mathbf{\Theta}_k \mathbf{W}\mathbf{C}_0\mathbf{W}^{-1}\mathbf{\Theta}_k^*,$$

where

$$\mathbf{\Theta}_k = diag([1, e^{-j2\pi k/N}, e^{-j2\pi 2k/N}, \dots, e^{-j2\pi Nk/N}])$$

and the * represents the complex conjugate. For the case of M=2,

$$\mathbf{W}\mathbf{C}_0\mathbf{W}^{-1} = \frac{1}{2}\begin{bmatrix} \mathbf{I}_{N/2} & \mathbf{I}_{N/2} \\ \mathbf{I}_{N/2} & \mathbf{I}_{N/2} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \otimes \mathbf{I}_{N/2}. \tag{6}$$

For other values of M, eq.(6) can be generalized. If we define the diagonal phase matrix as

$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{I}_N & 0 & 0 & 0 \\ 0 & \mathbf{\Theta}_1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \mathbf{\Theta}_{P-1} \end{bmatrix},$$

then $\tilde{\mathbf{C}}$ can be written

$$\tilde{\mathbf{C}} = \mathbf{\Phi}(\mathbf{I}_P \otimes \tilde{\mathbf{C}}_0)\mathbf{\Phi}^*$$

where the˜represents the Fourier transform of the spatial domain quantity.

Taking the Fourier transform of the restoration matrix in eq.(5), we have

$$\begin{aligned} \tilde{\mathbf{D}} &= \mathbf{K}_\lambda \otimes (\tilde{\mathbf{K}}_{xx}\tilde{\mathbf{H}}^T)\tilde{\mathbf{C}}^* \cdot \\ & [\tilde{\mathbf{C}}(\mathbf{K}_\lambda) \otimes (\tilde{\mathbf{H}}\tilde{\mathbf{K}}_{xx}\tilde{\mathbf{H}}^T) + \tilde{\mathbf{K}}_{\eta\eta})\tilde{\mathbf{C}}^*]^\dagger \end{aligned}$$

At this point, the computation of the solution follows the same path as that given in [4]. The bookkeeping is much simpler.

The computation, as shown in [4], is made tractable by using partitioning of the matrices. The details of that manipulation are not presented here. The major result is that the solution of an M-channel system with N samples requires only $M \times M$ matrix operations after taking the DFT of each of the channels. The simple form of the subsampling matrices and their DFTs makes finding the pseudoinverses easy.

4 Simulation

The method was simulated in MATLAB for the two-channel case. The baseline comparison is with the bandlimited reconstruction (MSE BL) obtained by assuming the signal is limited to the Nyquist frequency of the sampling matrix $\mathbf{C}_0$. The bandlimited reconstruction is obtained by treating each channel independently. This is the reconstruction that would be obtained if no knowledge of the other channels was available. The error of the estimate from our method is denoted MSE IL (interleaved).

From eq.(7), we see there are four parameters that will affect the quality of the solution: the impulse response of the sampling system, the temporal and between channel correlation of the signals and the correlation of the noise. The extent of the impulse response, $\mathbf{H}$, is determined by the circuitry that performs the actual sampling. While the ideal sampling would extend only to a single sample interval, practical considerations make the extent slightly larger. We have modelled this impulse response by an exponential, similar to the classic RC circuit. The time constant associated with this impulse ranges from 1 to 10 samples.

An important assumption that permits a tractable solution is the separability of the signal correlation into its temporal and between channel parts. The between channel correlation is given by an $M \times M$ matrix; the temporal correlation is given by an $N \times N$ matrix. For the simulations in this paper, the temporal correlation was obtained by passing white noise through a FIR filter, similar to the RC filter used for the system impulse response. The correlation of the signal is much longer, ranging from 8 to 16 samples.

The between channel correlation was modelled by using linear combinations of two uncorrelated signals. First we generate three temporally correlated signals, $s_1(t)$, $s_2(t)$ and $s_3(t)$, from uncorrelated, white Gaussian noise inputs with unit variance. The two channels are then generated by $x_1(t) = (1-\alpha)s_1(t)+\alpha s_2(t)$ and $x_2(t) = (1-\alpha)s_1(t-t_0)+\alpha s_3(t)$. The relative magnitude of α controls the between channel correlation. The parameter t_0 controls the time delay, so we can simulate the physical offset of two microphones that record the same signal.

The additive noise that simulates recording errors is uncorrelated with the M-channel signal and has a Gaussian distribution, $\mathbf{K}_{\eta\eta} = \sigma_\eta^2\mathbf{I}$.

To test the effect of bandwidth of the signals, we ran cases with the raw signals, which are assumed bandlimited to the Nyquist limit associated with the sampling rate, and with bandlimited signals that were filtered to have no power above the bandlimit of subsampled signals. For the bandlimited case, the additional channel correlation will improve the reconstruction only to the extent of overcoming the degradation caused by the additive noise.

5 Results

The results of a sampling of the experiments are summarized in Table 1. The columns of the table are describes below:

1. Indicator of the input signal being bandlimited (BL)or not (NBL), and an indicator of the method of computing the correlation matrix. EXCT indicates the correlation was computed using the actual simulated signal; APRX indicates the correlation was computed using the statistical model.

2. Signal-to-noise ratio in dB. Recall the noise is white Gaussian.

3. Between channel correlation coefficient. Since we are using a two channel system, this describes the correlation between channels.

4. Bandlimited Error measured in SNR with respect to the original input signal. There is a column for each channel.

5. Interleaved Error measured in SNR with respect to the original input signal. There is a column for each channel.

For this table, the time constant for the impulse response for the correlation of the signal is 16 samples; the time constant for the system impulse response is 3 samples. A complete test of the algorithm included variations of the these parameters, as well as the delay between channels and additional signal-to-noise ratios.

The qualitative results are as expected, which justifies the use of the abbreviated results presented in Table 1. The complete results are available on the web at http://www4.ncsu.edu/ hjt/. Comparing the first case of bandlimited versus nonbandlimited input with no noise, we see the advantage of using the second channel to help obtain higher resolution. The reconstruction based on the bandlimited assumption will treat the aliased portion of the signal as noise, whereas the interleaved method will treat it as information.

We tested the effect of the estimation of the correlation matrices by using estimates based on the exact signals generated for the simulations and based on the theoretical correlation derived from the signal generation model. The results show that using the approximate correlation causes little degradation in the performance.

6 Discussion

The demosaicking method based on baysian estimation has shown to produce noticeable improvements in the estimation of higher resolution sampling. Study of the method will be continued with the lifting of various constraints, such as the identical blur on each channel and integral spacing of the sampling.

References

[1] A. J. Patti, M. I. Sezan and A. M. Tekalp, "Superresolution Video Reconstruction with Arbitrary Sampling Lattices and Nonzero Aperture Time," IEEE Trans. Image Proc., Vol. 6, No. 8, pp. 1064-1176, August 1997.

[2] M. Elad and A. Feuer, "Restoration of a Sinfle Superresolution Image From Several Blurred, Noisy and Undersampled Measured Images," IEEE Trans. Image Proc., Vol. 6, No. 12, pp. 1646-1658, Dec. 1997.

[3] M. K. Ng, J. Koo and N. K. Bose, "Constrained Total Least-squares Computations for High-resolution Image Reconstruction with Multisensors," Int. J. of Imaging Systems and Technology, Vol. 12, No. 1, pp. 35-42, Dec. 2002.

[4] H. J. Trussell and R. E. Hartwig, "Mathematics for Demosaicking," IEEE Trans. Image Proc., Vol. 11, No. 4, pp. 485-492, April 2002.

Table 1: Results of Resolution Enhancement

BL/CorrEst	SNR	σ_{12}	BL Error SNR		IL Error SNR	
BL/EXCT	40	1	35.64	35.64	41.66	41.66
NBL/EXCT	40	1	9.53	9.39	37.04	37.04
NBL/APRX	40	1	9.53	9.39	37.04	37.04
BL/EXCT	40	0.87	35.64	35.64	39.76	37.86
NBL/EXCT	40	0.87	9.09	10.29	13.98	14.53
NBL/APRX	40	0.87	9.09	10.29	13.89	14.56
BL/EXCT	40	0.71	35.64	35.64	39.75	37.86
NBL/EXCT	40	0.71	8.72	10.62	11.88	12.70
NBL/APRX	40	0.71	8.72	10.62	11.63	12.78
BL/EXCT	40	0.5	35.64	35.64	39.75	37.86
NBL/EXCT	40	0.5	8.26	10.73	10.38	11.87
NBL/APRX	40	0.5	8.26	10.73	10.13	11.74
BL/EXCT	20	1	15.64	15.64	21.67	21.67
NBL/EXCT	20	1	9.14	8.00	17.13	17.13
NBL/APRX	20	1	9.14	8.00	17.11	17.11
BL/EXCT	20	0.87	15.64	15.64	19.94	18.01
NBL/EXCT	20	0.87	8.79	8.84	13.11	12.96
NBL/APRX	20	0.87	8.79	8.84	13.24	12.98
BL/EXCT	20	0.71	15.64	15.64	19.81	17.91
NBL/EXCT	20	0.71	8.44	9.20	11.48	11.79
NBL/APRX	20	0.71	8.44	9.20	11.40	11.80
BL/EXCT	20	0.5	15.64	15.64	19.78	17.88
NBL/EXCT	20	0.5	8.00	9.44	10.14	11.17
NBL/APRX	20	0.5	8.00	9.44	9.98	11.01