SPECTRALLY ADAPTED MERCER KERNELS FOR SUPPORT VECTOR SIGNAL INTERPOLATION

C. Figuera[1], *J. L. Rojo-Alvarez*[1], *M. Martinez-Ramon*[2], *A. Guerrero-Curieses*[1], *A. J. Caamaño*[1]

[1]Dept. of Signal Theory and Communications, Rey Juan Carlos University
Camino del Molino S/N, 28943, Fuenlabrada, Spain
[2] Dept. of Signal Theory and Communications, Carlos III de Madrid University,
Avda. de la Universidad 30, 28911, Leganés, Spain

ABSTRACT

Interpolation of nonuniformly sampled signals in the presence of noise is a hard and deeply analyzed problem. On the one hand, classical approaches like the Wiener filter use the second order statistics of the signal, and hence its spectrum, as *a priori* knowledge for finding the solution. On the other hand, Support Vector Machines (SVM) with Gaussian and *sinc* Mercer kernels have been previously proposed for time series interpolation, with good properties in terms of regularization and sparseness. Hence, in this paper we propose to use SVM-based algorithms with kernels having their spectra adapted to the signal spectrum, and to analyze their suitability for nonuniform interpolation. For this purpose, we investigate the performance of the SVM with autocorrelation kernels for one-dimensional time series interpolation. Simulations with synthetic signals show that SVM-based algorithms with the proposed kernels provide good performance for signals with different kinds of spectrum, even in the case of highly nonuniform sampling.

1. INTRODUCTION

The interpolation of nonuniformly sampled noisy signals is a very hard problem. A wide variety of solutions have been proposed by extending the original ideas by Shannon. A seminal work in this setting was presented by Yen [1], who proposed to use the *sinc* kernel and minimize the Mean Squared Error (MSE) of the reconstructed signal in the available observed samples. Other interpolation algorithms using the *sinc* kernel have been proposed [2], in which the *sinc* weights are obtained according to the minimization of the maximum error committed on the observed data set. In [3] the minimization of the sum of the MSE and a functional that penalizes the lack of smoothness is formulated for finding the best set of splines for the reconstruction of the signal. Also, different methods based on the Lagrange interpolator have been proposed for band-limited signal reconstruction [4, 5]. Recently, Support Vector Machines (SVM) algorithms for nonuniform signal interpolation have been proposed [6], using *sinc* and Gaussian kernels. In that work it was shown that the SVM algorithms provided very high performance for non-uniform signal interpolation. This is due to SVM regularization capabilites, robustness in the presence of different types of noise, sparse solutions and simplicity. SVM algorithms are formulated in terms of Mercer kernels, and a well-known theoretical result is that any autocorrelation function is a valid Mercer kernel [7]. However, to our best knowledge, no study has analyzed the suitability of using a Mercer kernel in SVM interpolation algorithms taking into account the spectral adaptation between the observed signal and the kernel.

In this paper we propose the use of SVM algorithms to solve the nonuniform sampling interpolation problem by exploring several Mercer kernels that are spectrally adapted to the original signal. For this purpose, we first analyze the relationship between the Wiener filter and the SVM for nonuniform interpolation going into the spectral interpretation of both algorithms. Then, according to this analysis, we study SVM algorithms whose kernels are spectrally adapted to the target signal, using for this purpose the signal autocorrelation.

This work was supported by the Spanish Ministry of Science and Innovation (MICINN) through projects TEC2009-12098 and TEC2010-19263

2. ALGORITHMS FOR NONUNIFORM INTERPOLATION

Problem statement. Let $x(t)$ be a continuous time signal, consisting of a signal $z(t)$ corrupted with noise, where the noise is modeled as a Wide Sense Stationary (WSS) process. This signal has been observed on a set of N unevenly spaced time instants, obtaining the observations $x = [x(t_1), \cdots, x(t_n), \cdots, x(t_N)]^T$. Then, the interpolation problem consists on finding a continuous-time signal $\hat{z}(t)$ which approximates $z(t)$ in another set of K different time instants, i.e., $z = [z(t'_1), \cdots, z(t'_k), \cdots, z(t'_K)]^T$, which is the unknown.

Wiener Filter for Interpolation. As described in [8], a Bayesian approach to solve this problem can be found by using the Wiener filter [9]. To obtain the Bayesian estimator $\hat{z}(t'_k)$, we apply the Linear Minimum Mean Square Error (LMMSE) estimator

$$\hat{z}(t'_k) = \sum_{n=1}^{N} a_{k,n} x(t_n) + a_{k,0} \quad (1)$$

for $k = 1, \ldots, K$, where coefficients $a_{k,n}$ and $a_{k,0}$ are chosen to minimize the mean square error. If x and z have zero mean, then $a_{k,0} = 0$, and the scalar LMMSE estimator is

$$\hat{z}(t'_k) = c^T_{z(t'_k)x} C^{-1}_{xx} x = a^T_k x \quad (2)$$

for $k = 1, \ldots, K$, where $c_{z(t'_k)x}$ represents the cross covariance column vector between the observed signal and the signal interpolated at time instant t'_k, and C_{xx} is the covariance matrix of the observations. After some manipulations, the coefficients for the Wiener filter are computed as

$$a_k = \left(R_{zz} + \sigma_w^2 I_N\right)^{-1} r_{zz}^{(k)} \quad (3)$$

where R_{zz} is the autocovariance matrix of the signal, σ_w^2 is the noise power, I_N is the identity matrix, and $r_{zz}^{(k)}$ is the cross covariance vector between the observed signal and the interpolated signal at instant t_k. Two main drawbacks arise when observing Eq. (2). On the one hand, there may exist more unknowns than observations, so the problem is ill conditioned; on the other hand, the autocorrelation of the signal is needed, so it must be estimated if it is not known.

Yen Regularized Interpolator. As an alternative to the Wiener filter for interpolation, *a priori* information can be used for band-limited signals, which is inspired in Shannon's Theorem for interpolation of noise-free signals. The signal model is stated as follows,

$$x(t'_k) = z(t'_k) + w(t'_k) = \sum_{n=1}^{N} a_n \mathrm{sinc}(\sigma_0(t'_k - t_n)) + w(t'_k) \quad (4)$$

where $sinc(t) = \frac{\sin(t)}{t}$, parameter $\sigma_0 = \frac{\pi}{T_0}$ is the *sinc* function bandwidth, and $w(t'_k)$ is the noise. Yen [1] proposed to use a Least Squares strategy to estimate coefficients a_n. Formally, when introducing a regularization term to prevent ill-posed problem, the vector $a = [a_1, \cdots, a_N]^T$ is given by

$$a = \left(S^2 + \delta I_N\right)^{-1} Sx \quad (5)$$

where the elements in S are computed as $S(k,n) = sinc(\sigma_0(t_k - t_n))$. Here, δ is a scalar parameter which controls the stability and smoothness of the solution, and which must be previously adjusted to completely define the algorithm.

SVM Interpolation. Another alternative is the SVM approach [6, 10]. Let us assume an interpolation model of the form

$$x(t) = v^T \varphi(t) \quad (6)$$

where v is a weight vector which defines the solution and $\varphi(t)$ is a nonlinear function in a Hilbert Space $\mathscr{H}$ provided with a dot product

$$\varphi(t_1)^T \varphi(t_2) = K(t_1, t_2) \quad (7)$$

where $K(\cdot,\cdot)$ is the kernel function, that must satisfy the Mercer's Theorem. The SVM criterion is intended to optimize the functional

$$\mathscr{L}_p = \frac{1}{2}\|v\|^2 + \frac{1}{2\gamma}\sum_{n\in I_1}(\xi_i^2 + \xi_i^{*2}) + C\sum_{n\in I_2}(\xi_i + \xi_i^*) \quad (8)$$

subject to the constraints

$$\begin{aligned} x_i - v^T\varphi(t_i) &\le \varepsilon - \xi_i \\ -x_i + v^T\varphi(t_i) &\le \varepsilon - \xi_i^* \end{aligned} \quad (9)$$

where $\{\xi_i\}$ and $\{\xi_i^*\}$ are the positive and negative slack variables and I_1 and I_2 are respectively the quadratic and linear sections of the cost function

$$\mathscr{L}^\varepsilon(e_n) = \begin{cases} 0, & |e_n| < \varepsilon \\ \frac{1}{2\gamma}(|e_n| - \varepsilon)^2, & \varepsilon \le |e_n| < e_C \\ C(|e_n| - \varepsilon) - \frac{1}{2}\gamma C^2, & |e_n| \ge e_C \end{cases} \quad (10)$$

where ε, γ and C are free parameters of the SVM algorithm and $e_C = \varepsilon + \gamma C$. A Lagrange optimization of functional (8) with constraints (9) and cost function (10) leads to the solution

$$v = \sum_{n=1}^{N}(\alpha_n - \alpha_n^*)\varphi(t_n) \quad (11)$$

where α, α^* are the Lagrange multipliers of Wolfe's dual [10]

$$\begin{aligned} \mathscr{L}_d = &-\frac{1}{2}(\alpha - \alpha^*)^T(R + \gamma\mathbf{I})(\alpha - \alpha^*) \\ &+ (\alpha - \alpha^*)^T x - (\alpha + \alpha^*)^T 1\varepsilon \end{aligned} \quad (12)$$

These multipliers must satisfy the constraints $0 \le \alpha_i, \alpha_i^* \le C$. Once these coefficients are found with a quadratic programming algorithm, by combining (6), (7) and (11), the dual formulation of the estimator is

$$x(t'_k) = \sum_{n=1}^{N}(\alpha_n - \alpha_n^*)K(t_n - t'_k) \quad (13)$$

The kernel can be seen as a time invariant system which provides a convolutional model for the solution [11].

3. SPECTRALLY ADAPTED MERCER KERNELS

In this section, a frequency domain interpretation of the Wiener and SVM interpolators is provided, and several Mercer kernels with different degrees of spectral adaptation are proposed for nonuniform sampling interpolation using SVM principles.

3.1. Spectral analysis for the algorithms

Wiener filter. The solution of the LMMSE estimator given by (2) can be seen as the convolution of the observations with a filter with impulse response $h_W^{(k)}[n] = a[k-n]$, which, in turn, can be used to provide a spectral interpretation of the estimator. Assuming that $N \to \infty$, it can be shown [8] that the transfer function of the filter is

$$H_W(f) = \frac{P_{zz}(f)}{P_{zz}(f) + P_{ww}(f)} = \frac{\eta(f)}{\eta(f) + 1} \quad (14)$$

where $P_{zz}(f)$ and $P_{ww}(f)$ are the power spectral density of the original signal and the noise respectively, and $\eta = \frac{P_{zz}(f)}{P_{ww}(f)}$ represents the local Signal to Noise Ratio (SNR) in a frequency f. Obviously, $0 \le H_W(f) \le 1$, tending to 1 in spectral bands with high SNR. The autocorrelation of the process to be interpolated is an indicator of the relevance of each spectral band in terms of SNR.

SVM algorithm. From the Karush-Kuhn-Tucker (KKT) conditions of the SVM interpolation algorithm [6], we can express the solution of the SVM interpolator in (13) as

$$\hat{z}(t) = \beta(t) * K(t) \quad (15)$$

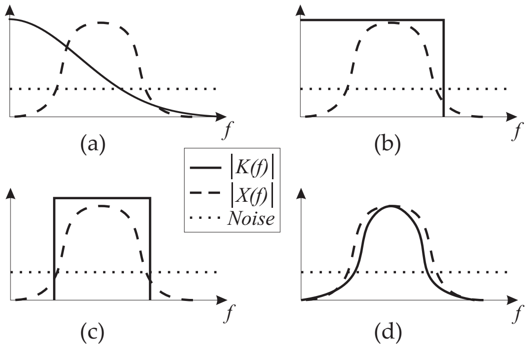

Fig. 1. Mercer kernels in the frequency domain ($|K(f)|$) and spectral adaptation to the band-pass signal ($|X(f)|$) in the presence of noise: non adapted low-pass RBF and Sinc kernels (a),(b), and adapted with modulated sinc and autocorrelation kernels (c),(d).

for a continuous time equivalent model of the equations and taking into account that we can express $\beta(t)$ as

$$\beta(t) = \sum_{k=1}^{K} (\alpha_k - \alpha_k^*) \delta(t - t_k) \tag{16}$$

with $\delta(t)$ the Dirac delta function.

The *sinc* or the Gaussian kernels represent low-pass transfer functions, and hence, their spectral adaptation to the observations can be poor when used for interpolation of a band-pass signal whose spectrum is not included into the kernel passband. However, in Eq. (14) we can see that the autocorrelation can be seen as a more accurate template in the frequency domain.

Nevertheless, despite these are well known properties in the signal processing literature, little attention has been paid to the use of spectrally adapted Mercer kernels for solving nonuniform interpolation problems. These are described next.

3.2. Mercer Kernels for SVM Interpolation

When examining the SVM signal model in (15), the role of the Mercer kernel is similar to a transfer function which recovers the interpolated signal from the linear system defined by the time series of the Lagrange multipliers. This suggests that Mercer kernels represent the transfer functions that emphasize the recovered signal in those bands with higher SNR. Hence we propose to use several Mercer kernels for signal interpolation with different degrees of adaptation to the spectral profile of the observations.

Inspired in the Wiener filter approach, we use the autocorrelation of the signal to be interpolated. Then, the following autocorrelation kernel can be defined

$$K(t_n, t'_k) = R(t_n - t'_k) \tag{17}$$

where R is the autocorrelation function of the process. Similarly to the Wiener filter case, since this autocorrelation will not always be available, an estimation procedure must be used. However, due to the robustness of the SVM algorithm, simple procedures for estimating the autocorrelation functions can be used. In this paper, a simple Lomb periodogram is used.

Figure 1 qualitatively shows the effect of using different kernels when interpolating a band-pass signal. In (a) and (b), the RBF and *sinc* kernels are used. Noise at low frequencies is undesirably amplified by the kernel. In (c) and (d) two band-pass adapted kernels (a modulated *sinc* and the autocorrelation of the signal) are used. In these cases the noise is properly filtered, which in turn enhances the result of the interpolation process.

4. EXPERIMENTS

For benchmarking the SVM with spectrally adapted kernels, we considered the following algorithms: the Yen algorithm with regularization (Yen); two versions of the Wiener filter, one with the autocorrelation estimated from the observations using a Lomb periodogram(Wien) and the other with the actual autocorrelation (Wien-Id); two SVMs with low-pass kernels (SVM-RBF, with Gaussian kernel, and SVM-Sinc), a SVM with a band-pass kernel constructed with a modulated *sinc* (SVM-ModSinc), and two SVMs, one with actual autocorrelation kernel (SVM-CorrId) and the other the kernel estimated with a Lomb periodogram (SVM-Corr). The central frequency and the bandwidth of the SVM-ModSinc algorithm are free parameters that should be chosen a priori.

A one-dimensional synthetic signal was interpolated by starting from a set of $L = 32$ unevenly spaced samples, with an average sampling interval $T = 0.5s$, and it was reconstructed by using a uniform grid with step $T_{int} = T/16$. The nonuniform sampling instants were simulated by adding a random quantity taken from a uniform distribution in the range $[-T/10, T/10]$ to the equally spaced time instants. The performance of each algorithm was measured with the S/E indicator, which is the ratio between the power of the signal and the power of the error in dB. Each experiment was repeated 50 times.

In order to show the importance of the spectral adaptation of the kernel, we apply all the algorithms described above but the SVM-ModSinc, when the original signal was a modulated squared *sinc* function, defined by

$$f(t) = sinc^2\left(\frac{\pi}{T_0}t\right)\cos\left(\frac{3\pi}{2}t\right) \tag{18}$$

Figure 2 shows an example of the spectra of the original and reconstructed signals and the error of reconstruction of each algorithm. When using the SVM algorithms with low-pass kernels, it can be observed that at low frequencies, where there is no significant signal power, the noise is enhanced, which produces a high error. Nevertheless, by observing the autocorrelation kernel spectrum, it can be seen that it is adapted to the signal spectrum and improves the filtering process, reducing the error at low frequencies. Figure 3 represents the performance of the algorithms for different SNR's. It can be observed that the two SVM interpolators with ideal and estimated autocorrelation kernels clearly outperform the rest of the algorithms (4-5 dBs and 2-3 dBs, respectively).

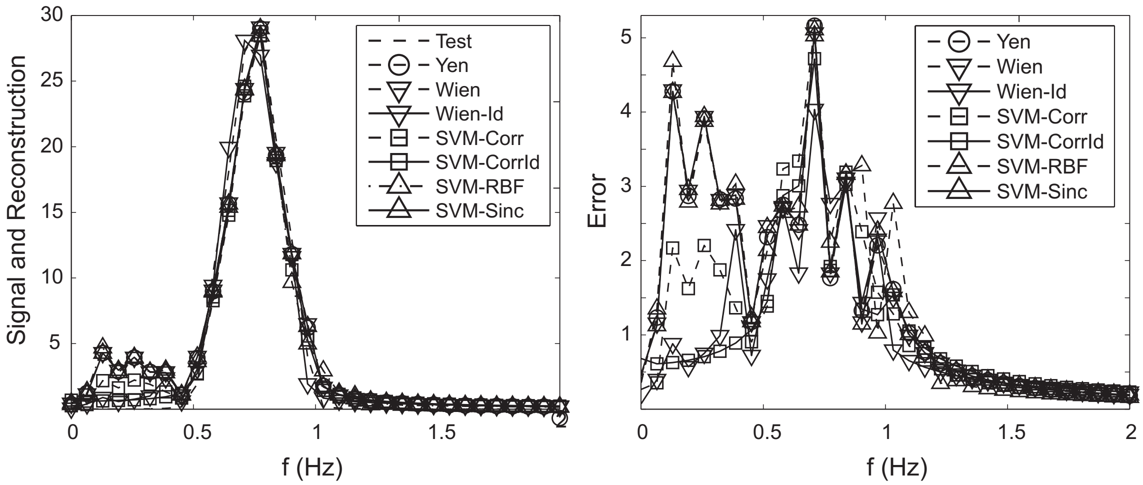

Fig. 2. Example of the spectra of the original and reconstructed signals (left) and the error of the reconstructions (right).

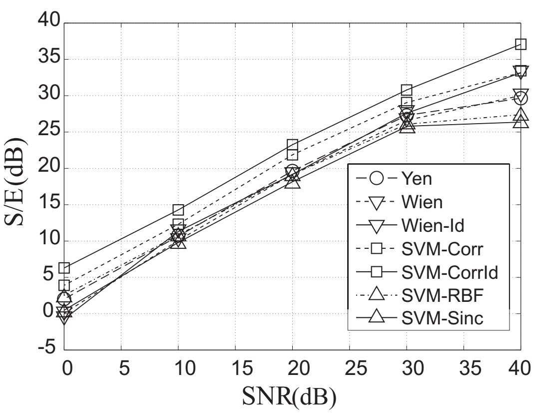

Fig. 3. S/E ratio against SNR for different algorithms for a band-pass interpolated signal.

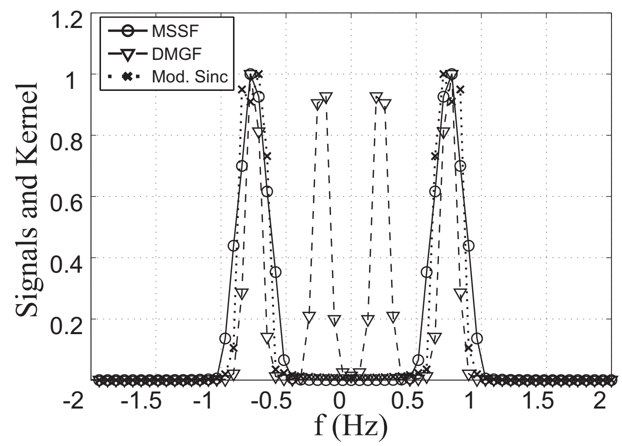

Fig. 4. Spectrum for the MSSF, DMGF and Modulated *Sinc* Kernel.

In order to emphasize the importance of the spectral matching of the kernel, we analyzed the performance of two SVM-based interpolators on two different test signals. The first signal is the modulated squared *sinc* function (MSSF) described in (18). The second signal consists of two Gaussian functions modulated at different frequencies and added together (called DMGF, see Figure 4). We interpolated these two functions by using the SVM-CorrId and the SVM-ModSinc. Typical Radial Basis Function (RBF) such as Gassuian kernel, and also *sinc* kernels can be adapted to band-pass signal interpolation by just being modulated by a sinusoidal function. The following is the modulated band-pass version (centered at ω_0) of the *sinc* kernel

$$K(t_n, t'_k) = \mathrm{sinc}(\sigma_0(t_n - t'_k)) \sin\left(\omega_0(t_n - t'_k)\right) \qquad (19)$$

The transfer function of this kernel is easily adapted to the spectral profile of the observed band-pass signal.

Figure 5 shows the S/E performance for both algorithms for the two functions. Since the autocorrelation kernel is able to adapt its spectrum, it performs well with both signals. However, the SVM-ModSinc performs well when applied to the MSSF, since the spectrum is similar (as can be observed in Figure 4) but the performance degrades when applied to the DMGF. This experiment highlights the relevance of the spectral adaptation of the kernel in SVM based interpolation problems.

Now, we examine the effect of nonuniform sampling over the algorithms performance. The sampling random quantity added to the uniform grid is now uniformly distributed in the range $[-u, u]$. Figure 6 shows the performance of six different interpolation algorithms for a set of values of u, from $u = 10^{-3}$ to $u = T/2 = 0.25$. With u very small the sampling is almost uniform, and with $u = T/2$ the samples can be placed at any time instant.The ideal autocorrelation kernel is more robust against the effect of the non-uniform sampling. When u takes its maximum value, the difference between SVM-CorrId and the rest of the algorithms rises up to 5.5 dB. Interestingly, the ideal Wiener filter behaves similar to the ideal autocorrelation kernel SVM for low values of u

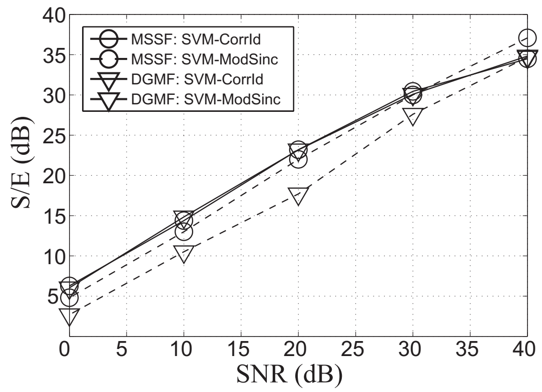

Fig. 5. S/E ratio for SVM-CorrId (solid lines) and SVM-ModSinc (dashed lines) algorithms when interpolating MSSF (circles) and DMGF (triangles).

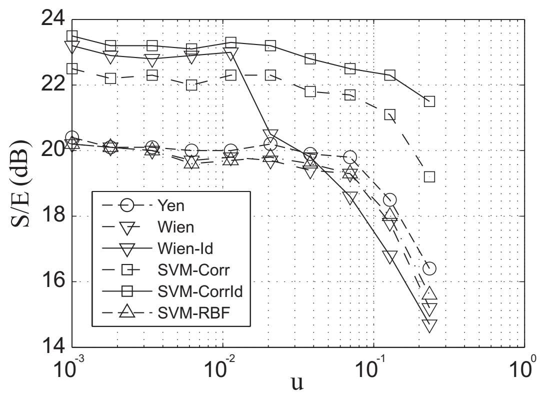

Fig. 6. S/E ratio for different values of the non-uniformity parameter (*u*).

which was the expected behavior, since both of them use the prior knowledge about the second order statistics of the signal. However, when u exceeds 0.1 the performance of the Wiener filter degrades fast, while the regularization properties of the SVM-based algorithms provides a higher robustness.

5. CONCLUSIONS

This paper presents two SVM algorithm which use spectrally adapted kernels for nonuniform interpolation. Although for this problem the SVM methodology provides a robust estimation of the signal, the spectral properties of the kernels and their impact in the performance of the interpolation algorithms had not yet been analyzed. For this purpose, we first examined the relationship between the Wiener filter and the SVM interpolation in the frequency domain. This allows us to motivate the search of a spectrally adapted kernel for the SVM algorithm. Then, we proposed to use the autocorrelation of the signal as the kernel for SVM interpolation, both the estimated autocorrelation and the ideal autocorrelation. The experiments show the capability of the autocorrelation kernels to adapt their spectra to the signal spectrum, and provide high performance even with highly nonuniform sampling. In the case when the autocorrelation of the signal is not known, a first estimation of the signal autocorrelation must be made. Note that for this purpose a simple interpolation algorithm (linear interpolation in our examples) can be used. Hence, a two-step process is needed for real signal: first a coarse estimation of the autocorrelation is made, and then the robustness of the SVM framework is used for an accurate estimation of the interpolated signal. We can conclude that both kernels can be useful kernels to be used for interpolation within the SVM signal processing framework.

6. REFERENCES

[1] J. L. Yen, "On nonuniform sampling of bandwidth-limited signals," *IRE Trans. Circuit Theory*, vol. 3, no. 4, pp. 251–57. 1956.

[2] H. Choi and D. C. Munson, "Direct-Fourier reconstruction in tomography and synthetic aperture radar," *Int'l. J. Imaging Systems and Technology*, vol. 9, no. 1, pp. 1–13. 1998.

[3] M. Arigovindan, M. Süthling, P. Hunziker, and M. Unser, "Variational image reconstruction from arbitrarily spaced samples: A fast multiresolution spline solution," *IEEE Trans. on Image Processing*, vol. 14, no. 4, pp. 450–460, 2005.

[4] J. Selva, "Functionally weighted lagrange interpolation of band-limited signals from nonuniform samples," *IEEE Trans. on Signal Processing*, vol. 57, no. 1, pp. 168–181, 2009.

[5] C. E. Shin, M. B. Lee, and K. S. Rim, "Nonuniform sampling of bandlimited functions," *IEEE Trans. on Information Theory*, vol. 54, no. 8, pp. 3814–3819, 2008.

[6] J. L. Rojo-Álvarez, C. Figuera-Pozuelo, C .E. Martínez-Cruz, G. Camps-Valls, F. Alonso-Atienza, and M. Martínez-Ramón, "Nonuniform Interpolation of Noisy Signals Using Support Vector Machines," *IEEE Trans. on Signal Processing*, vol. 55, no. 8, pp. 4116–4126, 2007.

[7] L Zhang, W Zhou, and L Jiao, "Wavelet support vector machine," *IEEE Trans Syst Man and Cybern-part B: Cybern*, vol. 34, no. 1, pp. 34–39, 2004.

[8] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, 1993.

[9] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, Wiley, New York, 1949.

[10] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 4, no. 3, pp. 199–222. 2004.

[11] J. L. Rojo-Álvarez, M. Martínez-Ramón, J. Muñoz-Marí, G. Camps Valls, C. M. Cruz, and A. Figueiras-Vidal, "Sparse deconvolution using Support Vector Machines," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, (2008).