# Design of linear-phase FIR filters using support vector regression approach

I. Santamaría

The design of linear-phase FIR filters is formulated within the support vector regression (SVR) framework. A simple design example shows that, for a given filter order, the proposed SVR technique yields slightly lower passband and stopband ripples than a minimax approach.

*Introduction:* The support vector (SV) method is a powerful learning technique for solving function estimation problems that has received considerable attention in recent years [1, 2]. By employing the so-called Vapnik's $\varepsilon$-insensitivity loss function, here we show that the design of linear-phase FIR filters can be formulated as a regression problem within the support vector machine (SVM) framework. To the best of our knowledge this is the first application of the powerful theory of SVMs to FIR filter design. The resulting cost function is convex with a global optimal solution that can be found through efficient quadratic programming (QP) algorithms [3]. A lowpass design is used to illustrate the performance of the proposed technique in comparison to a minimax design.

*FIR filter design via SVR:* For the sake of conciseness, we consider only the design of type I linear-phase FIR filters. Let $h[n]$ be the impulse response of a causal FIR filter with $P+1$ coefficients, where $P$ is the order of the filter. For a type I filter, $P=2M$ and $h[P-n]=h[n]$, for $0 \le n \le P$. Because of the even symmetry property, the frequency response of the filter is $H(\omega)=e^{-jM\omega}\bar{H}(\omega)$, where $\bar{H}(\omega)$ is the zero-phase frequency response, which can be expressed as

$$\bar{H}(\omega) = \mathbf{w}^T \mathbf{g}(\omega) \tag{1}$$

where $\mathbf{w}=(w[0],\ w[1], \ldots,\ w[M])^T$ and $g(\omega)=(1,\ \cos(\omega), \ldots, \cos(M\omega))^T$. The relationship between the linear regressor $\mathbf{w}$ and the filter coefficients is: $w[n]=h[M]$ if $n=0$ and $w[n]=2h[M-n]$ otherwise.

The design of an FIR filter amounts to finding the coefficients of the linear regressor $\mathbf{w}$ in (1) such that the frequency response of the resulting filter satisfies some requirements. For example, for a lowpass filter the specifications are usually given as

$$1-\varepsilon_p \le \bar{H}(\omega) \le 1+\varepsilon_p \quad \text{for } \omega \in [0, \omega_p]$$
$$-\varepsilon_s \le \bar{H}(\omega) \le \varepsilon_s \quad \text{for } \omega \in [\omega_s, \pi]$$

where $\omega_p$ and $\omega_s$ are the passband and stopband edge frequencies, respectively; and the maximum admissible errors are $\varepsilon_p$ (passband) and $\varepsilon_s$ (stopband).

A number of techniques to obtain the 'best' FIR filter according to different approximation criteria exist in the literature [4]. Typically, least-squared or minimax error designs are used in practice. Here we consider the use of a SV-based technique: according to the structural risk minimisation (SRM) principle [1], to design a linear-phase FIR filter one minimises

$$J(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2 + C\left(\sum_{i\in[0,\omega_p]} |1-\mathbf{w}^T\mathbf{g}(\omega_i)|_{\varepsilon_p} + \sum_{i\in[\omega_s,\pi]} |\mathbf{w}^T\mathbf{g}(\omega_i)|_{\varepsilon_s}\right) \tag{2}$$

where

$$|y-\mathbf{w}^T\mathbf{g}(\omega_i)|_\varepsilon = \max\{0, |y-\mathbf{w}^T\mathbf{g}(\omega_i)|-\varepsilon\}$$

is the so-called Vapnik's $\varepsilon$-insensitive loss function and $C>0$ is a penalty value, which establishes a tradeoff between model complexity (minimising the squared norm of the filter coefficients $\|\mathbf{w}\|^2$), and the penalty given to deviations larger than $\varepsilon_p$ or $\varepsilon_s$.

The procedure starts by selecting an initial set of $N$ input frequencies $(\omega_1, \ldots, \omega_N)$ equispaced in the interval $\Omega=[0, \omega_p]\cup[\omega_s, \pi]$. For each frequency, an input–output pattern is constructed as $(\mathbf{g}(\omega_i), y_i)$, where

$$y_i = \begin{cases} 1 & \text{if } \omega_i \in [0, \omega_p] \\ 0 & \text{if } \omega_i \in [\omega_s, \pi] \end{cases} \tag{3}$$

and $\mathbf{g}(\omega_i)=(1, \cos(\omega_i), \ldots, \cos(M\omega_i))^T$.

From the dataset constructed in this way the SVM for regression can readily be derived. In particular, by introducing a set of positive slack variables $\xi_i$ and $\tilde{\xi}_i$, the minimisation of (2) can be rewritten as the following constrained optimisation problem: to minimise

$$L(\mathbf{w}, \xi, \tilde{\xi}) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}(\xi_i + \tilde{\xi}_i)$$

subject to

$$\mathbf{w}^T\mathbf{g}(\omega_i) - y_i \le \varepsilon_i + \xi_i \tag{4}$$
$$y_i - \mathbf{w}^T\mathbf{g}(\omega_i) \le \varepsilon_i + \tilde{\xi}_i \tag{5}$$
$$\xi_i, \tilde{\xi}_i \ge 0 \tag{6}$$

for all $i=1, \ldots, N$; where $y_i$ is given by (3) and $\varepsilon_i \in \{\varepsilon_p, \varepsilon_s\}$ denote the maximum permissible deviation in each band. Using the Lagrange multiplier technique [1], it is easy to derive that the optimal regressor (i.e. the optimal FIR filter) can be expanded as

$$\mathbf{w} = \sum_{i=1}^{N}(\tilde{\alpha}_i - \alpha_i)\mathbf{g}(\omega_i) \tag{7}$$

where $\tilde{\alpha}_i$ and $\alpha_i$ are positive Lagrange multipliers, which are obtained by maximising the following quadratic form (dual problem)

$$W(\alpha, \tilde{\alpha}) = -\sum_{i=1}^{N}\varepsilon_i(\tilde{\alpha}_i + \alpha_i) + \sum_{i=1}^{N}(\tilde{\alpha}_i - \alpha_i) - \frac{1}{2}\sum_{i,j=1}^{N}(\tilde{\alpha}_i - \alpha_i)(\tilde{\alpha}_j - \alpha_j)\langle\mathbf{g}(\omega_i), \mathbf{g}(\omega_j)\rangle$$

subject to $0 \le \alpha_i,\ \tilde{\alpha}_i \le C$; and where $\langle\mathbf{g}(\omega_i), \mathbf{g}(\omega_j)\rangle$ denotes the inner product between the input patterns. This is a convex quadratic programming (QP) problem and, therefore, has a globally optimal solution that can be efficiently solved [2, 3].

*Design example:* To illustrate the proposed procedure, we consider the design of a lowpass filter of order $P=50$ $(M=25)$ with passband $[0, 0.4\pi]$ and stopband $[0.5\pi, \pi]$. The maximum deviations between the actual and desired responses are $\varepsilon_p=0.01$ for the passband and $\varepsilon_s=0.001$ (60 dB) for the stopband. For this example we used $N=200$ equispaced frequencies in the interval $\Omega=[0, 0.4\pi]\cup[0.5\pi, \pi]$. The only parameter to be selected in advance in the SV regression is the regularisation parameter $C$. This value is sometimes chosen as the range of the training dataset [5], therefore we decided to use $C=1$. The resulting QP problem has been solved using the Matlab SVM toolbox available in [3].

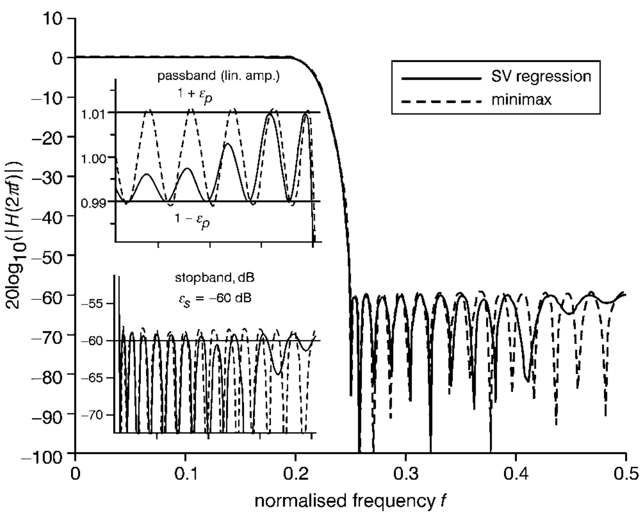


**Fig. 1** *Amplitude response of lowpass filter of order P = 50 designed with proposed SV regression procedure and with Remez algorithm*

Maximum error deviations are $\varepsilon_p=0.01$ for passband and $\varepsilon_s=0.001$ (60 dB) for stopband

For comparison an optimal filter of the same order has been designed using the Remez multiple exchange algorithm (minimax) [4]. Fig. 1

shows the frequency responses of the lowpass FIR filters designed using the proposed SV regression technique and the Remez algorithm. We can see that both design procedures provide similar filters. Nevertheless, a more detailed plot reveals some differences: the passband and stopband ripples for SVR-based FIR filter always fall below the maximum allowable deviation, whereas the FIR designed with Remez has higher peak ripples (only 59 dB of stopband attenuation). The same behaviour has been found in a number of examples.

From the initial set of $N = 200$ input patterns, only 26 became support vectors for this example, i.e. patterns for which exactly one of the Lagrange multipliers in (7) is greater than zero. Fig. 2 shows the SV coefficients ($\tilde{\alpha}_i - \alpha_i$ in the expansion (7)) for each input pattern. Not surprisingly, the largest SVs tend to concentrate in the transition bandwidth.

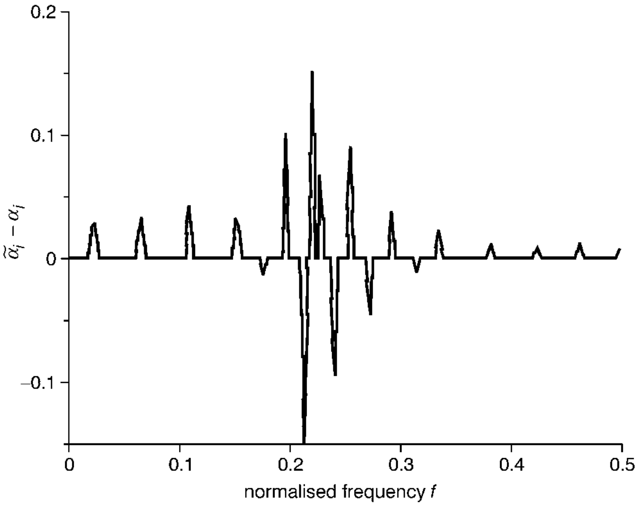


**Fig. 2** *Coefficients of SV expansion*

*Conclusions:* In this Letter the design of linear-phase FIR filters using the SVM methodology has been studied. A simple design example has shown that some interesting differences exist with respect to a minimax design. Despite its computational cost, the proposed technique has some potential advantages: for instance, it can be easily extended to incorporate additional constraints in the frequency or time domain. Moreover, similar ideas based on the SVM framework can be used for the design of nonlinear filters (by performing the linear regression in another space of higher dimension by the so-called kernel trick [2]).

*Acknowledgment:* This work was supported by MCYT (Ministerio de Ciencia y Tecnología) under grant TIC2001-0751-C04-03.

 *22 July 2003*
*Electronics Letters Online No: 20030914*
*DOI: 10.1049/el:20030914*

I. Santamaría (*DICOM, ETSII y Telecom., Universidad de Cantabria, Av. Los Castros s.n., 39005 Santander, Spain*)

E-mail: nacho@gtas.dicom.unican.es

## References

1 VAPNIK, V.: 'The nature of statistical learning theory' (Springer Verlag, New York, 1995)
2 SCHÖLKOPF, B., and SMOLA, A.: 'Learning with kernels' (The MIT Press, Cambridge, MA, 2002)
3 GUNN, S.R.: 'MATLAB support vector machine toolbox', Image Speech and Intelligent Systems Research Group, University of Southampton, UK, 1998 (available at http://www.isis.ecs.soton.ac.uk/isystems/kernel/)
4 SARAMÄKI, T.: 'Finite impulse response filter design' *in* MITRA, S.K., KAISER, J.F. (Eds.): 'Handbook for digital signal processing' (John Wiley & Sons, New York, 1993), Chap. 4, pp. 155–277
5 MATTERA, D., and HAYKIN, S.: 'Support vector regression for dynamic reconstruction of a chaotic system' *in* SCHÖLKOPF, B., BURGES, C.J.C., SMOLA, A.J. (Eds.): 'Advances in kernel methods: support vector learning' (The MIT Press, Cambridge, MA, 1999), pp. 211–241