TOWARDS OPTIMAL LEAST SQUARE FILTERS USING THE EIGENFILTER APPROACH

Cha Zhang and Tsuhan Chen

Dept. of Electrical and Computer Engineering, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213, USA
{czhang, tsuhan}@andrew.cmu.edu

ABSTRACT

In this paper, we propose a new eigenfilter approach to designing least square error filters. The filters are obtained by finding an eigenvector of a real, symmetric and positive definite matrix, which is numerically stable. The proposed algorithm has two advantages. First, we show that the least-square solution, which can only be obtained through matrix inversion in the literature, can be asymptotically reached with our algorithm. Second, when numerical errors break the matrix inversion method, our algorithm can still find some "optimal" filter through tuning an internal parameter.

1. INTRODUCTION

It is well know that most linear-phase finite impulse response filter (FIR) design problems can be solved by the McClellan-Parks (MP) algorithm [1]. The MP algorithm provides the optimal filter design results in the sense of minimizing the maximum error in both passband and stopband. On the other hand, a number of researchers have also considered the least-squares approach to FIR filter design [2]–[6]. Under certain situations the least-squares approach is preferable, such as when time- and frequency-domain constrains need to be incorporated.

There are two well-documented least-squares approaches to FIR filter designs: the matrix inversion (MI) method and the eigenfilter approach. The MI method is based on solving a set of linear equations by matrix inversion [2], and the eigenfilter approach is based on the computation of an eigenvector of an appropriate real, symmetric, and positive-definite matrix. Although the MI method tries to minimize the real square-error between the designed filter and the target filter, the process of MI is numerically unstable because we have to inverse a huge matrix with very small determinant. The inversion of the matrix is also computational expensive. The eigenfilter approach overcomes these problems by modifying the target function they try to minimize. For example, the first eigen-approach by Vaidyanathan and Nguyen (V&N) [4] added to the minimization problem a constraint that at the reference frequency, the designed filter has to have the same response as the target filter. Pei and Tseng (P&T) [6] removed the above constraint by using the total least squares (TLS) error criterion. In both approaches the filter design is formulated as the computation of an eigenvector of a real, symmetric, and positive-definite matrix, which is much more stable than the MI method. However, the price is that for many filter design problems, the square errors given by the eigenfilter approach is greater than the MI method.

In this paper, we propose a new framework to design least square error filters. In our framework, the filters are still obtained by finding an eigenvector of a real, symmetric and positive definite matrix, which keeps numerical stability. The P&T algorithm turns out to be a special case of our framework, while the least-square solution, which can only be obtained through MI method in the literature, can be asymptotically reached. The V&N algorithm is very similar to another extreme case of our algorithm, which will be discussed in the paper. When numerical errors break the MI method, with our algorithm we are able to find some "optimal" filter simply by tuning a parameter in our framework.

The paper is organized as follows. Section 2 gives a brief overview of existing least-squares approaches and proposes the new framework. Section 3 gives a filter design example where the MI method, the V&N algorithm, the P&T algorithm and our algorithm are compared. Section 4 shows another example where numerical errors play a role. Section 5 concludes the paper.

2. THE PROPOSED FRAMEWORK

2.1. Overview of previous approaches

Consider a typical filter design problem. We want to approximate a "desired response" $D(\omega)$ with a certain real-coefficient linear-phase FIR filter $H(\omega)$. We can rewrite $H(\omega)$ as the inner product of two vectors [3]-[6], i.e.:

$$H(\omega)=\mathbf{a}^T\mathbf{C}(\omega) \tag{1}$$

where $\mathbf{a}$ is a vector related with the coefficients of the designed filter, and $\mathbf{C}(\omega)$ is a vector of appropriate trigonometrical functions. The goal of least-squares approaches is to find the vector $\mathbf{a}$ that can minimize the square error between the target and the designed filters. In other words, the objective function of the minimization is:

$$J_{\mathrm{MI}}(\mathbf{a})=\int_{\omega\in R}\left[D(\omega)-\mathbf{a}^T\mathbf{C}(\omega)\right]^2 d\omega \tag{2}$$

where R is the region we care about, e.g., the passband and the stopband. The MI method solves the problem directly [2]:

$$\mathbf{a}=\mathbf{Q}_{\mathrm{MI}}{}^{-1}\mathbf{p} \tag{3}$$

where

$$\mathbf{Q}_{\mathrm{MI}}=\int_{\omega\in R}\mathbf{C}(\omega)\mathbf{C}^T(\omega)d\omega \tag{4}$$

and

$$\mathbf{p}=\int_{\omega\in R}D(\omega)\mathbf{C}(\omega)d\omega \tag{5}$$

Unfortunately, the inversion of matrix $\mathbf{Q}_{\mathrm{MI}}$ takes a long time and might be numerically unstable, which causes bad design results when the length of the filter is long.

Vaidyanathan and Nguyen [4] proposed to change the representation of $D(\omega)$ by adding a reference frequency ω_0, i.e.:

$$D'(\omega)=\frac{D(\omega)}{D(\omega_0)}\mathbf{a}^T\mathbf{C}(\omega_0) \tag{6}$$

Work supported by NSF Career Award 9984858.

which allows the target filter amplitude to be scaled by the frequency response of the designed filter at the reference point. On the other hand, it requires a normalization step to force the reference point frequency response to be equal to that of the original target filter $D(\omega)$. By making $\mathbf{a}^T\mathbf{a}=1$ at the same time to avoid trivial solutions, the objective function is changed to:

$$J_{\text{V\&N}}(\mathbf{a})=\frac{\mathbf{a}^T\mathbf{Q}_{\text{V\&N}}\mathbf{a}}{\mathbf{a}^T\mathbf{a}} \tag{7}$$

where:

$$\mathbf{Q}_{\text{V\&N}}=\int_{\omega\in R}\left[\frac{D(\omega)}{D(\omega_0)}\mathbf{C}(\omega_0)-\mathbf{C}(\omega)\right]\cdot\left[\frac{D(\omega)}{D(\omega_0)}\mathbf{C}(\omega_0)-\mathbf{C}(\omega)\right]^T d\omega \tag{8}$$

The nice thing is that the objective function (7) can be minimized by simply looking for the eigenvector corresponding to the minimum eigenvalue of the real, symmetric and positive-definite matrix $\mathbf{Q}_{\text{V\&N}}$, which is both computationally efficient and stable. However, there are two shortcomings of this algorithm. First, the designed filter is sensitive to the selection of the reference frequency, as will be shown in Section 3. Second, the constraint $\mathbf{a}^T\mathbf{a}=1$ is restricting and the above algorithm cannot achieve as small square error as the MI method when there is no numerical problem.

The more recent work by Pei and Tseng [6] improves the V&N algorithm by using the total least square (TLS) error criterion. They modify the objective function as:

$$J_{\text{P\&T}}(\mathbf{a})=\int_{\omega\in R}\frac{\left[D(\omega)-\mathbf{a}^T\mathbf{C}(\omega)\right]^2}{1+\mathbf{a}^T\mathbf{a}}d\omega \tag{9}$$

which is recognized as a traditional plane fitting problem minimizing the square distances of points $(\mathbf{C}(\varphi),D(\varphi)),\forall\varphi\in R$ and the superplane $D(\omega)-\mathbf{a}^T\mathbf{C}(\omega)=0$, where R is the frequency region we care about and $\mathbf{a}$ is related to the optimal filter coefficients we try to find. By defining:

$$\hat{\mathbf{a}}=\left[\mathbf{a}^T\quad -1\right]^T,\ \hat{\mathbf{C}}(\omega)=\left[\mathbf{C}(\omega)^T\quad D(\omega)\right]^T \tag{10}$$

We have:

$$J_{\text{P\&T}}(\mathbf{a})=\frac{\hat{\mathbf{a}}^T\mathbf{Q}_{\text{P\&T}}\hat{\mathbf{a}}}{\hat{\mathbf{a}}^T\hat{\mathbf{a}}} \tag{11}$$

where

$$\mathbf{Q}_{\text{P\&T}}=\int_{\omega\in R}\hat{\mathbf{C}}(\omega)\hat{\mathbf{C}}^T(\omega)d\omega \tag{12}$$

The optimal filter in the TLS sense is obtained by finding the eigenvector with respect to the minimum eigenvalue of $\mathbf{Q}_{\text{P\&T}}$. The authors show that by relaxing the reference frequency constraint and the $\mathbf{a}^T\mathbf{a}=1$ constraint, the resulting filter has a smaller square error than the V&N algorithm.

2.2. The unifying framework

We propose a new objective function for least-squares filter design as follows:

$$J_{\text{ours}}(\mathbf{a})=\int_{\omega\in R}\frac{\left[D(\omega)-\mathbf{a}^T\mathbf{C}(\omega)\right]^2}{\beta^2+\mathbf{a}^T\mathbf{a}}d\omega \tag{13}$$

where $\beta\in(0,+\infty)$ is a parameter we can tune. Notice that the P&T algorithm has a very similar objective function as ours. In fact, the P&T algorithm is simply a special case of our framework when $\beta=1$.

However, our objective function has much more implications than the P&T algorithm. Let us first look at one extreme case of β, i.e., $\beta\to+\infty$. The objective function changes to:

$$J_{\text{ours}}(\mathbf{a})\approx\int_{\omega\in R}\frac{\left[D(\omega)-\mathbf{a}^T\mathbf{C}(\omega)\right]^2}{\beta^2}d\omega \tag{14}$$

Since β is a pre-fixed constant during the design, we are getting the optimal least-squares filter which could only be achieved by the MI method in the literature!

On the other hand, when $\beta\to 0$, the objective function becomes:

$$J_{\text{ours}}(\mathbf{a})\approx\int_{\omega\in R}\frac{\left[D(\omega)-\mathbf{a}^T\mathbf{C}(\omega)\right]^2}{\mathbf{a}^T\mathbf{a}}d\omega \tag{15}$$

which is very similar to the objective function of the V&N approach. Both of them impose the constraint that $\mathbf{a}^T\mathbf{a}=1$. However, Equation (15) does not have the reference point constraint. On one hand, we do not need to worry about choosing the right reference point any more, which eases the filter design process. On the other hand, Equation (15) loses the flexibility of changing the target $D(\omega)$'s amplitude with $\mathbf{a}$ as in Equation (6), which hurts the filter design performance. As will be shown in Section 3, if we choose the best reference frequency for the V&N algorithm, the resulting filter is better than minimizing Equation (15). If the reference frequency is randomly chosen, minimizing Equation (15) will provides a smaller overall square error in general.

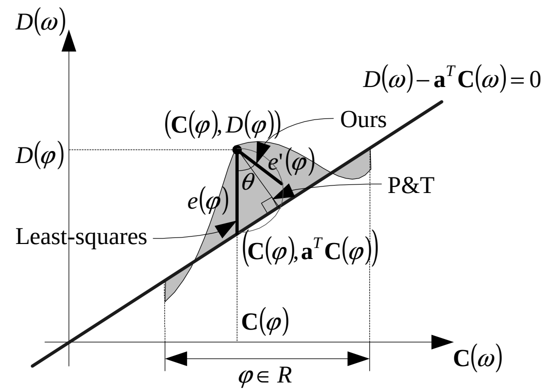

Figure 1 Geometric explanation of the unified framework.

The geometric explanation of the proposed framework is shown in Figure 1. The horizontal axis is $\mathbf{C}(\omega)$, which actually represents a vector of all the trigonometrical base functions; the vertical axis is $D(\omega)$. For the surface $(\mathbf{C}(\varphi),D(\varphi)),\forall\varphi\in R$ (shown as a curve in Figure 1), which is predetermined by the target filter and the base functions, we try to fit a plane $D(\omega)-\mathbf{a}^T\mathbf{C}(\omega)=0$ to it. The goal of least-squares filter design is, by adjusting the vector $\mathbf{a}$, to minimize the overall vertical distances of the surface to the plane. The objective function is:

$$J(\mathbf{a})=\int_{\varphi\in R}e^2(\varphi)d\varphi \tag{16}$$

where $e(\varphi)=D(\varphi)-\mathbf{a}^T\mathbf{C}(\varphi)$ is shown in Figure 1. However, the direct answer to Equation (16) is the MI method, which is expensive and sometimes numerically unstable. In our framework, instead of Equation (16), we minimize:

$$J_{\text{ours}}(\mathbf{a})=\int_{\varphi\in R}e'^2(\varphi)d\varphi=\int_{\varphi\in R}\left[e(\varphi)\cos\theta\right]^2d\varphi \tag{17}$$

where

$$\theta = \tan^{-1}\left(\frac{\sqrt{\mathbf{a}^T\mathbf{a}}}{\beta}\right) \quad (18)$$

This is equivalent to projecting the vertical error $e(\varphi)$ onto some line at a certain angle θ with respect to the vertical line and minimizing the projected error $e'(\varphi)$. In this way, we transfer the matrix inversion problem into an eigenfilter problem that is more numerically stable. Moreover, by increasing β, we decrease the θ angle and get closer and closer to the target of Equation (16), which is optimal. Again we can see that the P&T algorithm is the special case when $\beta = 1$ and $e'(\varphi)$ becomes the perpendicular distance from the point to the plane.

To minimize Equation (13), we can take an approach similar to the P&T algorithm. Define:

$$\hat{\mathbf{a}} = \begin{bmatrix}\mathbf{a}^T & -\beta\end{bmatrix}^T, \ \hat{\mathbf{C}}(\omega) = \begin{bmatrix}\mathbf{C}(\omega)^T & \frac{D(\omega)}{\beta}\end{bmatrix}^T \quad (19)$$

We have:

$$J_{\text{ours}}(\mathbf{a}) = \frac{\hat{\mathbf{a}}^T \mathbf{Q}_{\text{ours}} \hat{\mathbf{a}}}{\hat{\mathbf{a}}^T \hat{\mathbf{a}}} \quad (20)$$

where

$$\mathbf{Q}_{\text{ours}} = \int_{\omega \in R} \hat{\mathbf{C}}(\omega)\hat{\mathbf{C}}^T(\omega) d\omega \quad (21)$$

All we need to do is to find the eigenvector corresponding to the minimum eigenvalue of a real, symmetric and positive-definite matrix $\mathbf{Q}_{\text{ours}}$, and scale the eigenvector so that the last element of $\hat{\mathbf{a}}$ is $-\beta$. The filter coefficients can be found by mapping the $\mathbf{a}$ vector back according to how we obtain Equation (1). One thing to notice is that when β is very large or very small, using Equation (19) will introduce another source of numerical errors. Fortunately, experiment shows that we can already be very close to the optimal least-squares filter when β is around 10 to 100.

3. EXAMPLE I – A COMPARISON

In this section, we give an example to verify the ideas mentioned above. Assume that we are asked to design a Type-1 linear-phase low-pass FIR filter with least-squares approaches. The desired frequency response is given as:

$$D(\omega) = \begin{cases} 1, & 0 \le \omega \le \omega_p, \omega_p = 0.2\pi \\ 0, & \omega_s \le \omega \le \pi, \omega_s = 0.3\pi \\ \text{don't care} & \omega_p \le \omega \le \omega_s \end{cases} \quad (22)$$

The length of the filter is $N = 15$.

3.1. The MI solution

Since the order of the desired filter is low, the inversion of the matrix $\mathbf{Q}_{\text{MI}}$ is possible. The least square error filter can be solved directly through Equation (3). The overall square error

$$E^2 = \int_{\omega \in R} \left[D(\omega) - \mathbf{a}^T\mathbf{C}(\omega)\right]^2 d\omega \quad (23)$$

can be calculated as $E_{opt}^2 = 0.0077138$. If we use square error as the measure of the performance of a filter design algorithm, this is the best result we can ever achieve.

3.2. The V&N solution

With the V&N algorithm, we need to carefully choose the reference frequency in order to get the best performance.

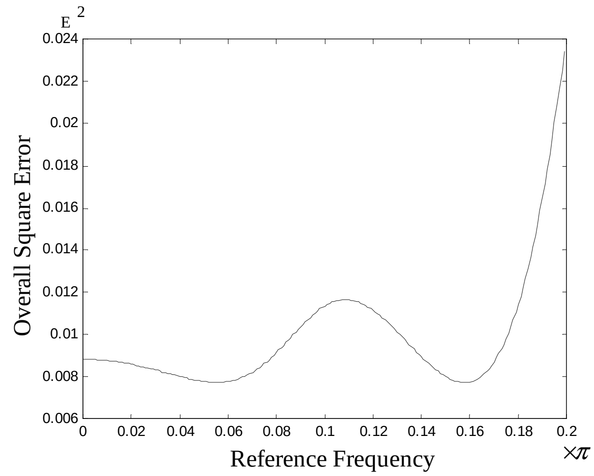

Figure 2 V&N algorithm with different reference frequency.

As shown in Figure 2, the overall square error of the resultant filter by the V&N algorithm changes when we choose different reference frequencies. The best result that can be reached is met when the reference frequency is 0.158π. The corresponding overall square error is $E^2 = 0.0077368$. We define *Distance To the Optimal (DTO)* as:

$$DTO = 10\log_{10}\left(\frac{E^2 - E_{opt}^2}{E_{opt}^2}\right)(dB) \quad (24)$$

where E^2 is the total square error of the current algorithm calculated by Equation (23), and E_{opt}^2 is the optimal square error that is obtained by the MI method. Obviously, the lower the DTO, the better the filter design algorithm. The V&N algorithm can reach a minimum DTO of –25.2516dB.

3.3. The solution of the proposed algorithm

In the proposed framework, we may change the value of β to get different filters. We plot the relationship between the DTO and β as in Figure 3. We can see that, in the worst case, when $\beta \to 0$, the DTO converges to about –21.5563dB, which is slightly worse than the best filter the V&N algorithm can obtain. However, if the reference frequency is randomly chosen between 0 and 0.2π in the V&N algorithm, 95% of the chance the V&N algorithm will give a worse performance than the worst case of our proposed framework. As $\beta \to +\infty$, we observe a very rapid drop of the DTO of the designed filter. This simply means that we can obtain the optimal least square error filter by increasing β. We see that when $\beta = 10$, $\text{DTO} = -69.5756\text{dB}$, which means we are very close to the least-squares solution already. Note that when $\beta > 10$, we plot the curve as dash line since the difference of E^2 and E_{opt}^2 in Equation (24) has already been smaller than the precision of the numerical integration we use to calculate them. However, the fast dropping tendency is still observable when $\beta > 10$.

4. EXAMPLE II – THE NUMERICAL STABILITY

It is very interesting to see the behavior of our algorithm under the condition of numerical errors. In this section, we design a group of filters with passband $[0, 0.2\pi]$ and stopband $[0.4\pi, \pi]$. The lengths of the filters are 15, 55, 95, 135 and 161. Since the MI method may suffer from numerical errors, we use $10\log_{10} E^2$ (E^2 as defined in Equation (23)) as the performance measure.

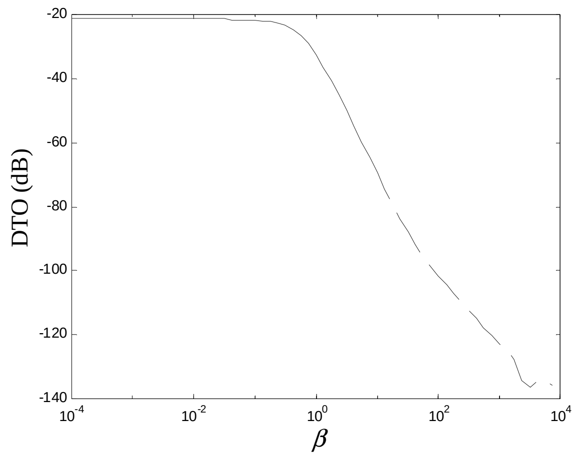

Figure 3 Choose different β in our algorithm.

We list the filter design results of all the approaches in Table I. For simplicity, we choose $\omega = 0$ as the reference frequency of the V&N algorithm. It is obvious that when the filter length is 15, 55 or 95, P&T and ours are better than V&N. In all the cases, our algorithm is always better than the P&T, as the former is a generalization of the latter. For the filters with length 135 and 161, where we can say that the filter design is highly numerical sensitive, the MI method breaks down. Note that in this range, the V&N actually outperforms or is similar to P&T and our algorithm. This is surprising but generally true based on some other experiments. The reason might be owing to the flexibility of adjusting the target filter's amplitude during the minimization (Equation (6)).

Table I: Filter design results using least-squares approaches $10\log_{10} E^2$ (dB)

	N=15	N=55	N=95	N=135	N=161
MI method	-28.7078	-90.9738	-82.8651	31.9093	29.7836
V&N	-28.0651	-90.7416	-150.3832	-137.6438	-129.5422
P&T	-28.7066	-90.9738	-150.5183	-110.1596	-119.2163
Ours	-28.7078	-90.9738	-150.5184	-122.9683	-126.4135

We show the relationship between the overall square error and the parameter β in Figure 4. We can draw a list of interesting conclusions from Figure 4. First, when the length of the filter is 15 or 55, the MI method is till applicable. Therefore, we see that for a wide range of β, the square errors keep the same. In fact, if we enlarge the figure, we will be able to see the monotonic decreasing of error when β increases. This is consistent with the experiment we had in Section 3. Second, when the filter length increases to 95, the MI method suffer a little bit from numerical errors, while our algorithm produces a extremely good result as long as $\beta > 0.1$. In the range $[10^{-1}, 10^4]$, we notice that the errors are gently fluctuating around –150dB, which we believe is caused by all the numerical errors during the whole process. Third, the results when the filter length is huge (N=135 and N=161) seem very interesting. They warn us that β cannot be arbitrary large. Introducing a very large β (or very small β) into the minimization problem may generate another source of numerical error. Fourth, there exists an optimal β for a certain filter design problem when numerical errors happen. This can be shown by verifying that both the curve of the filter with length 135 and 161 in Figure 4 have a convex shape. Fifth, we were expecting that there might be some horizontal shifting between the filters with length 135 and 161, so that different length of filter requires different optimal β. However, our experiments do not agree with this. On the contrary, $\beta \in [0.1, 1]$ seems to be a magic region that most probably we will find the optimal β in this region. Finally, increasing the length of the filter does not always mean smaller error, especially when numerical errors have to be considered. The best filter within the four is the one with length 95. Designing filters with too much length will not give good results due to the numerical error for all the algorithms we tested.

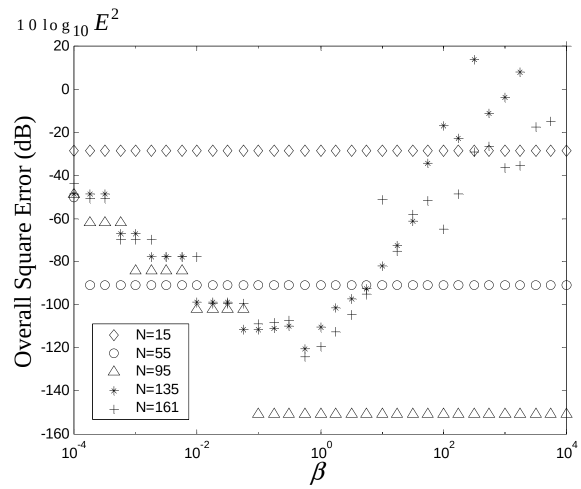

Figure 4 Behavior of our algorithm under numerical errors.

5. CONCLUSIONS

In this paper, we proposed a new eigenfilter based algorithm on least-square error filter design. We showed that we could asymptotically reach the optimal least-square error filter without matrix inversion. When numerical errors happen, we could find the optimal filter by simply tuning a parameter in our algorithm. The proposed algorithm can be easily modified to design filters with linear constraints, equal ripple filters, and multi-dimensional filters.

REFERENCES

[1] J. H. McClellan and T. Parks, "A Unified Approach to The Design of Optimum FIR Linear-Phase Digital Filters", *IEEE Trans. Circuit Theory*, vol. CT-20, pp. 697-701, Nov. 1973.

[2] D. W. Tufts and J. T. Francis, "Designing Digital Low-Pass Filters: Comparison of Some Methods and Criteria", *IEEE Trans. Audio Electroacoust.*, vol. AU-18, pp. 487-494, Dec. 1970.

[3] C. S. Burrus, A. W. Soewito, and R. A. Gopinath, "Least squared error FIR filter design with transition bands," *IEEE Trans. Signal Processing*, vol. 40, pp. 1327–1340, June 1992.

[4] P. P. Vaidyanathan and T. Q. Nguyen, "Eigenfilter: A new approach to least-squares FIR filter design and applications including Nyquist filters," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 11–23, Jan. 1987.

[5] T. Chen, "Unified Eigenfilter Approach: With Applications to Spectral/Spatial Filtering", *IEEE International Symposium on Circuits and Systems (ISCAS '93)*, May 1993.

[6] S. C. Pei and C. C. Tseng, "A New Eigenfilter Based on Total Least Squares Error Criterion", *IEEE Trans. Circuits Syst. I: Fundamental Theory and Applications*, vol. 48, pp. 699-709, June 2001.