PERFORMANCE MEASURES AND LAGRANGE MULTIPLIER METHODS TO TWO-BAND PR LP FILTER BANK DESIGN

Tao Wang and Benjamin W. Wah

Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
{wangtao,wah}@manip.crhc.uiuc.edu
URL: http://manip.crhc.uiuc.edu

ABSTRACT

In this paper, we study performance measures for designing two-band perfect reconstruction (PR) linear phase (LP) filter bank. Based on these performance metrics, we formulate the design problem as a nonlinear constrained optimization problem, where some metrics such as stopband energy have closed form, but the others like transition width do not. Our formulation allows us to search for designs that improve over the existing designs [6]. More important, given user-specified performance bounds such as maximal transition width, we are able to design filter banks if solutions exist, and trade-off among different performance metrics can be easily achieved. Finally, many experimental results show feasibility and efficiency of our filter bank design method.

1. INTRODUCTION

Digital filter banks [5] have been used in many engineering fields and applications such as audio and image coding. Their major advantage in processing signals and images is that they constitute a multirate information system.

There are two major approaches to design filter banks. In optimization-based methods, the design problem has been formulated as a multi-objective nonlinear optimization problem, whose form can be application- and filter-dependent, and then converted into a single-objective optimization problem and solved by existing optimization methods, such as gradient descent, Lagrange multiplier, quasi-Newton, simulated annealing, and genetics based methods. On the other hand, filter-bank design problems can also be solved using nonoptimization based algorithms, which include spectral factorization and heuristic methods. These methods generally do not continue to find better designs once a suboptimal design has been found.

Research supported in part by National Science Foundation Grants MIP 92-18715 and MIP 96-32316.

In this paper, we study a two-band FIR filter bank, which consists of an analysis system followed by a synthesis system. The analysis filters $H_0(z)$ and $H_1(z)$ decompose input signal $X(z)$, and the synthesis filters $G_0(z)$ and $G_1(z)$ reconstructs output $\hat{X}(z)$ from the subband signals. Choosing $G_0(z) = 2H_1(z)$ and $G_1(z) = -2H_0(-z)$ to cancel out aliasing term related to $X(-z)$, we obtain the relationship between input $X(z)$ and $\hat{X}(z)$ as follows

$$\hat{X}(z) = [H_0(z)H_1(-z) - H_0(-z)H_1(z)]X(z) \qquad (1)$$

Therefore, filter bank design problem becomes estimate of two filter parameters $h_0(n)$ and $h_1(n)$ that satisfy some performance metrics.

2. PERFORMANCE MEASURES

In this section, we identify filter bank design objectives and show how to formulate and evaluate them. The major contribution is that we are able to accurately measure some performances such as transition width, which is not studied in existing work [5]. Since all these performance metrics are involved in our optimization process, given user-specified performance bounds, we can find solutions which are impossible for existing work.

The performance metrics of the filter bank have to two parts. The first part relates to the overall filter bank response, and the second to the individual filters

2.1. Performance metrics for overall filter bank

To achieve both PR and LP properties in the filter bank, it is required that the sum of the filter lengths is a multiple of 4, i.e, $N_0 + N_1 = 4k$, where N_0 and N_1 are lengths of filters $H_0(z)$ and $H_1(z)$, respectively, and then the system delay of the filter bank is $(N_0 + N_1)/2 - 1$.

Only two types of nontrivial filter bank systems have both PR and LP features [4]. As illustration in this paper, we only discuss the case where both filters $H_0(z)$ and $H_1(z)$ have even length, $H_0(z)$ is symmetric, and $H_1(z)$ is anti-symmetric. Hence, two filters need to be estimated. One is low-pass filter $H_0(z)$ with parameters $h_0 = \{h_0(n), n = 0, 1, \cdots, N_0/2 - 1\}$, and the other is high-pass filer $H_1(z)$ with parameters $h_1 = \{h_1(n), n = 0, 1, \cdots, N_1/2 - 1\}$

The PR condition [1] can be enforced by a set of equality constraints,

$$\frac{1}{2}\theta\left(i - \frac{N_0 + N_1}{2}\right) = \sum_{k=0}^{2i-1}(-1)^k h_0(2i - 1 - k)h_1(k) \quad (2)$$

where $i = 1, 2, \cdots, \frac{N_0+N_1}{4}$, and $\theta(x) = 1$ if $x = 0$ and 0 otherwise.

2.2. Performance metrics for individual filters

Performance metrics for individual filters include stopband energy $E_s(h_0)$ and $E_s(h_1)$, stopband ripple $\delta_s(h_0)$ and $\delta_s(h_1)$, passband energy $E_p(h_0)$ and $E_p(h_1)$, passband ripple $\delta_p(h_0)$ and $\delta_p(h_1)$, and transition width $\tau(h_0)$ and $\tau(h_1)$. Figure 1 illustrates their definitions.

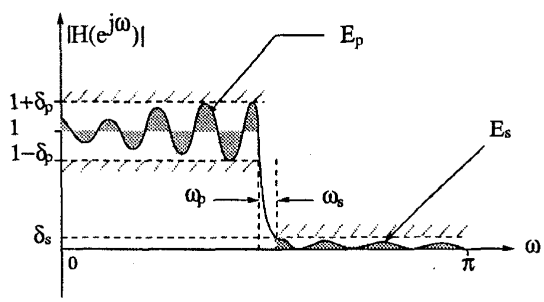

Figure 1: Performance metrics of individual filters.

Note that Figure 1 refers to a low-pass filter. For a high-pass filter $H_1(z)$, we use the same method to calculate its performance metrics on its mirror $H_1(-z)$. Hence, it is enough to illustrate our computation on low-pass filter H_0.

For given filter parameters h_0, both stopband and passband energy $E_s(h_0)$ and $E_p(h_0)$ depend on cut-off frequency $\omega_s(h_0)$ and $\omega_p(h_0)$, which vary with filter parameters h_0. This means that existing work that uses predefined values ω_s and ω_p does not give precise performance metrics, and thus it can not ensure satisfaction of user-defined performance bounds.

In addition, stopband and passband ripples $\delta_s(h_0)$ and $\delta_p(h_0)$, and transition width $\tau(h_0)$ do not have closed form equations. We have to evaluate all of them using numerical methods.

The stopband ripple $\delta_s(h_0)$ is estimated in the frequency range $[\pi/2, \pi]$ using the absolute value $|H_0(\omega)|$ of frequency response $H_0(\omega)$. First, we uniformly sample the frequency range from $\pi/2$ to π. Frequency intervals that contain a local maxima are found based on the sample points. Within each interval, we apply Newton's method to precisely locate the local maximum points. When Newton's method does not converge after a fixed number of iterations, Golden search is used to reduce the interval and re-start Newton's method. Among all these local maxima, the largest one is the stopband ripple $\delta_s(h_0)$. In a similar way, the passband ripple $\delta_p(h_0)$ is found in frequency range of $[0, \pi/2]$.

Using the stopband ripple $\delta_s(h_0)$, we are able to find stopband cut-off frequency $\omega_s(h_0)$, which is the first $\omega \in [\pi/2, \pi]$ such that $|H_0(\omega)| = \delta_s(h_0)$. This is done by first finding $\omega_s(h_0)$ using sampling, and then refining it with combination of Newton's method and bisection search. Similarly, the passband cut-off frequency $\omega_p(h_0)$ is the first $\omega \in [0, \pi/2]$ such that $|H_0(\omega)| = 1 - \delta_p(h_0)$.

With estimated cut-off frequencies $\omega_s(h_0)$ and $\omega_p(h_0)$, we can easily calculate transition width $\tau(h_0) = \omega_s(h_0) - \omega_p(h_0)$, stopband energy $E_s(h_0)$, as well as passband energy $E_p(h_0)$. Note that, when we calculate these performance metrics, the main operation that may cause lower precision is the sum in $|H_0(\omega)|$. In order to maintain at high precision, we employ Kahan's summation formula.

3. CONSTRAINED OPTIMIZATION FORMULATION

Obviously, designing PR LP filter bank has multiple objectives for both overall filter bank and individual filters: (a) satisfy PR condition (2), i.e. equality constraints; (b) minimize stopband energy $E_s(h_0)$ and $E_s(h_1)$; (c) minimize stopband ripple $\delta_s(h_0)$ and $\delta_s(h_1)$; (d) minimize passband energy $E_p(h_0)$ and $E_p(h_1)$; (e) minimize passband ripple $\delta_p(h_0)$ and $\delta_p(h_1)$; (f) minimize transition width $\tau(h_0)$ and $\tau(h_1)$.

It can be formulated either as a multi-objective unconstrained optimization or as a single-objective constrained optimization. In a multi-objective formulation, a possible way is to optimize the design with respect to a subset of the performance measures, for example,

$$\text{PR condition} \quad \text{and} \quad \min\ E_s(h_0) \text{ and } E_s(h_1) \quad (3)$$

which is adopted by most existing work [5]. Unfortunately, optimal solutions to the simplified optimization problem are not necessarily optimal solutions to the original problem, and performance measures not included in the formulation are compromised.

In general, optimal solutions of a multi-objective problem form a *Pareto optimal frontier* such that one solution on this frontier is not dominated by another. One approach to find a point on the Pareto frontier is to optimize a weighted sum of all the objectives. This approach has difficulty when Pareto frontier points of certain characteristics are desired, such as those with certain transition width. Different combinations of weights must be tested by trial and error until a desired filter is found. When the desired characteristics are difficult to satisfy, trial and error is not effective in finding feasible designs. In this case, constrained formulation should be used instead.

In constrained formulation, constraints are defined with respect to a reference design or user-specified performance bounds. Constraint-based methods have been applied to design filter banks in both the frequency [1] and the time domains [3]. In all these designs, most performance metrics like transition width and ripples are not evaluated, and energy using predefined cut-off frequencies is not accurate. Therefore, given a performance bound, none of them can be used to design a required filter bank.

We formulate the filter bank design in the most general form as a constrained nonlinear optimization problem, in the sense that all the performance metrics are evaluated precisely based on their definitions, as follows,

$$\begin{array}{lll} \textit{Minimize} & E_{PR}(h_0, h_1) & \\ \textit{Subject to} & \text{PR condition} & \\ & E_s(h_0) \le \tilde{E}_{s0} & E_s(h_1) \le \tilde{E}_{s1} \\ & E_p(h_0) \le \tilde{E}_{p0} & E_p(h_1) \le \tilde{E}_{p1} \\ & \delta_s(h_0) \le \tilde{\delta}_{s0} & \delta_s(h_1) \le \tilde{\delta}_{s1} \\ & \delta_p(h_0) \le \tilde{\delta}_{p0} & \delta_p(h_1) \le \tilde{\delta}_{p1} \\ & \tau(h_0) \le \tilde{\tau}_0 & \tau(h_1) \le \tilde{\tau}_1 \end{array} \tag{4}$$

where $\tilde{E}_{s0}, \tilde{E}_{s1}, \tilde{E}_{p0}, \tilde{E}_{p1}, \tilde{\delta}_{s0}, \tilde{\delta}_{s1}, \tilde{\delta}_{p0}, \tilde{\delta}_{p1}, \tilde{\tau}_0$, and $\tilde{\tau}_1$ are user-specified performance bounds. The objective function $E_{PR}(h_0, h_1)$ is defined as a squared sum of all the equality constraints in PR condition (2). Hence, the minimum of the objective function is equivalent to PR condition.

The performance metrics cover different ranges of values, for example, the stopband energy is of order 10^{-4}, the ripples of order 10^{-2}, and the transition width of order 10^{-1}. In order to balance their effects on optimization, we choose to normalize them and obtain

$$\begin{array}{lll} \textit{Minimize} & E_{PR}(h_0, h_1) & \\ \textit{Subject to} & \text{PR condition} & \\ & E_s(h_0)/\tilde{E}_{s0} - 1 \le 0 & E_s(h_1)/\tilde{E}_{s1} - 1 \le 0 \\ & E_p(h_0)/\tilde{E}_{p0} - 1 \le 0 & E_p(h_1)/\tilde{E}_{p1} - 1 \le 0 \\ & \delta_s(h_0)/\tilde{\delta}_{s0} - 1 \le 0 & \delta_s(h_1)/\tilde{\delta}_{s1} - 1 \le 0 \\ & \delta_p(h_0)/\tilde{\delta}_{p0} - 1 \le 0 & \delta_p(h_1)/\tilde{\delta}_{p1} - 1 \le 0 \\ & \tau(h_0)/\tilde{\tau}_0 - 1 \le 0 & \tau(h_1)/\tilde{\tau}_1 - 1 \le 0 \end{array} \tag{5}$$

If the performance bounds are computed from the best known design and we tighten some of them, it is possible for us to improve the solution. For a simpler case of QMF filter bank design, we have already shown some improved solutions and efficiency of our algorithm [6].

4. LAGRANGE MULTIPLIER METHODS

To solve the constrained optimization problem (5), we use a Lagrange multiplier method [2], where inequality constraints are transformed into equality constraints using slack variables. In the Lagrangian formulation, a local minimum in a feasible region is a *saddle point* at which the objective function is at a local minimum and the weighted sum of the constraints is at a local maximum. By using this property, saddle points can be found by local search methods that perform gradient descents in the original-variable space and gradient ascents in the Lagrange-variable space.

The Lagrange multiplier algorithm requires first-order derivatives to compute gradients. For closed-form formulas such as stopband energy $E_s(h_0)$, it is easy to derive their analytical forms of derivatives. But for nonclosed-form formula such as transition width $\tau(h_0)$, we use finite difference methods to approximate their derivatives.

Because the formulated filter bank problem is highly nonlinear constrained optimization problem, the Lagrange multiplier method may get trapped into a local minimum like most local strategies such as gradient descent and Newton's method. To escape from the local minimum, global search approaches can be utilized, for instance, covering methods, interval methods, simulated annealing and genetic algorithms. But they take much longer time than local search methods, and thus tradeoff between quality of solution and computational time has to be examined.

The goal of this paper is to show how our formulation and evaluation can be used to design the filter bank given user-specified performance bounds and how to do trade-off among different performance metrics. We will not study performance of different global search algorithms.

5. EXPERIMENTAL RESULTS

Starting from an initial point $\left(h_0^{(t=0)}, h_1^{(t=0)}\right)$, we set initial values of the Lagrange multipliers to be zero, and solve the dynamic equations of the Lagrangian function by using *LSODE* [1] until it converges.

Here, we only describe one example to show how our formulation improves existing results and makes tradeoff among different performance metrics. We use example 2.3 [1] as our baseline where $N_0 = 16$ and $N_1 = 28$.

To design such filters, the authors [1] use the fixed frequency cut-offs, i.e. $w_s(h_0) = 0.6\pi$, $w_p(h_0) = 0.44\pi$, $w_s(h_1) = 0.6\pi$, and $w_p(h_1) = 0.4\pi$. However, according to definition of the frequency cut-off, these values should vary with the filter parameters. Even for the reported solution, the actual cut-offs should be $w_s(h_0) = 0.623\pi$, $w_p(h_0) = 0.419\pi$, $w_s(h_1) = 0.620\pi$, and $w_p(h_1) = 0.465\pi$. This

[1] *LSODE* is a solver for first-order ordinary differential equations, a public-domain package available from http://www.netlib.org.

Table 1: Improved result for 95% passband ripple.

performance of filter h_0	normalized solution	performance of filter h_1	normalized solution
$E_s(h_0)$	1.00	$E_s(h_1)$	1.00
$E_p(h_0)$	0.96	$E_p(h_1)$	0.96
$\delta_s(h_0)$	1.00	$\delta_s(h_1)$	1.00
$\delta_p(h_0)$	0.95	$\delta_p(h_1)$	0.95
$\tau(h_0)$	1.00	$\tau(h_1)$	1.00

Table 2: Improved result for 90% passband ripple.

performance of filter h_0	normalized solution	performance of filter h_1	normalized solution
$E_s(h_0)$	1.00	$E_s(h_1)$	0.98
$E_p(h_0)$	1.00	$E_p(h_1)$	1.00
$\delta_s(h_0)$	1.00	$\delta_s(h_1)$	1.00
$\delta_p(h_0)$	0.90	$\delta_p(h_1)$	0.90
$\tau(h_0)$	1.00	$\tau(h_1)$	1.00

Table 3: Trade-off for stopband ripple and passband energy.

performance of filter h_0	normalized solution	performance of filter h_1	normalized solution
$E_s(h_0)$	0.99	$E_s(h_1)$	0.81
$E_p(h_0)$	1.05	$E_p(h_1)$	1.04
$\delta_s(h_0)$	0.90	$\delta_s(h_1)$	0.90
$\delta_p(h_0)$	1.00	$\delta_p(h_1)$	1.00
$\tau(h_0)$	1.00	$\tau(h_1)$	1.00

Table 4: Trade-off for reduced transition width.

performance of filter h_0	normalized solution	performance of filter h_1	normalized solution
$E_s(h_0)$	1.00	$E_s(h_1)$	0.87
$E_p(h_0)$	1.74	$E_p(h_1)$	1.69
$\delta_s(h_0)$	1.00	$\delta_s(h_1)$	1.00
$\delta_p(h_0)$	1.00	$\delta_p(h_1)$	1.00
$\tau(h_0)$	0.95	$\tau(h_1)$	0.94

means performance metrics of stopband and passband energy are not accurately calculated. In addition, performance metrics of ripples and transition widths are not measured.

Using the solution given by [1], we first compute performance metrics, and set them as our performance bounds $\tilde{E}_{s0}$, $\tilde{E}_{s1}$, $\tilde{E}_{p0}$, $\tilde{E}_{p1}$, $\tilde{\delta}_{s0}$, $\tilde{\delta}_{s1}$, $\tilde{\delta}_{p0}$, $\tilde{\delta}_{p1}$, $\tilde{\tau}_0$, and $\tilde{\tau}_1$ in our constrained formulation (5). In the following, our solutions are normalized by these bounds, meaning that the value less than one is better.

In the first set of experiments, we tighten performance bounds of passband ripples to show improved results. To obtain the design with 95% of the passband ripples of both filters h_0 and h_1, we set performance bounds as $0.95\tilde{\delta}_{p0}$ and $0.95\tilde{\delta}_{p1}$, and solve it using the Lagrange multiplier method. All the constraints are satisfied, and the performance metrics of our solution is given in Table 1. We improve both passband ripples and energy. The performance bounds of passband ripples are further tightened to 90%, and we find a solution shown in Table 2.

In the next set of experiments, we want to study how to do trade-off among different performance metrics, where some metrics are relaxed while some are tightened. Table 3 shows the result that we tighten stopband ripples to 90% and relax passband energy by 5%.

Table 4 shows another trade-off result for reduced transition width. Obviously, transition width is a critical performance metric. Tightening it little bit may cause large degradation of other performance metrics like passband energy.

As a conclusion, two points have to be emphasized. First, the reason why the desired filter bank can be designed given performance bounds is that we accurately evaluate these performance metrics, and formulate it as a constrained optimization problem. Second, involving all these metrics makes the design problem more difficult to solve due to high nonlinear and non closed form constraints. Therefore, the solutions we obtained are just local search results, which can be further improved using some global search method.

6. REFERENCES

[1] B. R. Horng and A. N. Willson Jr. Lagrange multiplier approaches to the design of two-channel perfect-reconstruction linear-phase fir filter banks. *IEEE Trans. on Signal Processing*, 40(2):364–374, February 1992.

[2] D. G. Luenberger. *Linear and Nonlinear Programming*. Addison-Wesley Publishing Company, 1984.

[3] K. Nayebi, T. P. Barnwell III, and M. J. T. Smith. Time-domain filter bank analysis: A new design theory. *IEEE Transactions on Signal Processing*, 40(6):1412–1429, June 1992.

[4] T. Q. Nguyen and P. P. Vaidyanathan. Two-channel perfect reconstruction fir qmf structure which yield linear-phase analysis and synthesis filters. *IEEE Trans. on Acoutics, Speech, and Signal Processing*, 37(5):676–690, May 1989.

[5] P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice Hall, Englewood Cliffs, New Jersey, 1993.

[6] B. W. Wah, Y. Shang, T. Wang, and T. Yu. QMF filter bank design by a new global optimization method. In *Proc. Midwest Symposium on Circuits and Systems*, Iowa City, Iowa, August 1996. IEEE.