Optimum Design of Balanced SAW Filters Using Evolutionary Multi-objective Optimization

KIYOHARU TAGAWA
School of Science and Engineering
Kinki University
3-4-1 Kowakae, Higashi-Osaka, 577-8502
JAPAN
tagawa@info.kindai.ac.jp

Abstract: The frequency response characteristics of the balanced Surface Acoustic Wave (SAW) filters are governed primarily by their geometrical structures. Therefore, in order to realize desirable frequency response characteristics, the structural design of the balanced SAW filter is formulated as a constrained multi-objective optimization problem. Then a recent Evolutionary Multi-objective Optimization (EMO) method, which is called Generalized Differential Evolution 3 (GED3), is applied to the multi-objective optimization problem. Furthermore, in order to clarify the tradeoff relationship among the objective functions of the multi-objective optimization problem, Principal Component Analysis (PCA) is used to assess the set of the non-dominated solutions obtained by GDE3.

Key–Words: Surface acoustic wave filter, evolutionary algorithm, optimum design

1 Introduction

Surface Acoustic Wave (SAW) filters are small, rugged and cost-competitive mechanical band-pass filters with outstanding frequency response characteristics. Therefore, SAW filters have played an important role as a key device in various mobile and wireless communication systems such as personal digital assistants (PDAs) and cellular phones[1, 2]. Recently, the balanced SAW filter becomes widely used in the modern Radio Frequency (RF) circuits of cellular phones. That is because the balanced SAW filter can provide not only the band-pass filtering function but also some external functions such as the unbalance-balance signal conversion, the impedance conversion and so on[3]. Consequently, by using the balanced SAW filter, we can reduce the total number of the components of the modern RF circuit, as well as their mounted area. As a result, we can miniaturize the modern RF circuits of cellular phones.

The frequency response characteristics of SAW filters including balanced ones are governed primarily by their geometrical structures, namely, the configurations of Inter-Digital Transducers (IDTs) and Shorted Metal Strip Arrays (SMSAs) reflectors fabricated on piezoelectric substrates. Therefore, in order to realize desirable frequency response characteristics of SAW filters, we have to decide their suitable structures, which are specified by some design parameters such as the numbers of the electrodes of IDTs.

In order to decide a suitable structure of the SAW filter, optimum design methods that combine the optimization algorithm with the computer simulation have been reported. In some of them, Evolutionary Algorithms (EAs) such as Genetic Algorithm (GA) have been also used as the optimization algorithm[4, 5]. In our previous paper[5], a recent EA called Differential Evolution (DE)[6] was applied to the optimum design problem of a practical balanced SAW filter.

Specifications for the balanced SAW filter are described by using several criteria. However, in the previous optimum design method[5], the structural design of the balanced SAW filter was formulated as a single-objective optimization problem. Exactly speaking, a single-objective function was defined by the weighted sum of the several criteria of the balanced SAW filter. One difficulty in the previous optimum design method is the choice of appropriate weighting coefficients. Therefore, even if a very good solution could be obtained for the single-objective optimization problem, we do not necessarily obtain a desirable structure of the balanced SAW filter.

In this paper, a multi-objective optimum design method for balanced SAW filters is proposed. First of all, the structural design of the balanced SAW filter is formulated as a constrained multi-objective optimization problem. In the multi-objective optimization problem, three objective functions about the attenuation of the balanced SAW filter are defined respectively within three different bandwidths. Furthermore, besides the boundary constraints on the design parame-
etters, six non-linear constraints are considered.

In order to obtain various Pareto-optimal solutions for the above multi-objective optimization problem, a recent Evolutionary Multi-objective Optimization (EMO) method, which is called Generalized Differential Evolution 3 (GED3)[7], is employed. GDE3 is an extension of DE for global optimization with an arbitrary number of objectives and constraints over continuous space. However, the design parameters of the balanced SAW filter take not only continuous values but also discrete values. Therefore, in order to apply GED3 to the structural design of the balanced SAW filter, we employ a technique that represents various design parameters by using only real-parameters.

In order to demonstrate the usefulness of the proposed multi-objective optimum design method, GDE3 is applied to a three-objective optimum design problem of a practical balanced SAW filter. Furthermore, Principal Component Analysis (PCA) is used to assess the set of the non-dominated solutions obtained by GDE3. As a result, it is shown that two of the three objective functions are in the tradeoff relationship.

2 Balanced SAW Filter

2.1 Structure and Principle

The balanced SAW filter consists of several components, namely, Inter-Digital Transducers (IDTs) and Shorted Metal Strip Array (SMSA) reflectors fabricated on a piezoelectric substrate. Figure 1 illustrates a typical structure of the balanced SAW filter that consists of nine components: one transmitter IDT (IDT-T), two receiver IDTs (IDT-Rs), pitch-modulated IDTs between IDT-T and IDT-R, and two SMSA reflectors. In the balanced SAW filter shown in Fig. 1, port-1 is an unbalanced input-port, while a pair of port-2 and port-3 is a balanced output-port.

2.2 Modeling and Simulation

The network model of the balanced SAW filter shown in Fig. 1 can be represented graphically as shown in Fig. 2. In the network model, nodes \( a_q \) \((q = 1, 2, 3)\) denote the input signals of the balanced SAW filter, while nodes \( b_p \) \((p = 1, 2, 3)\) denote the output signals. Scattering parameters \( s_{pq} \) labeled on edges provide the transition characteristics between input signals \( a_q \) and output signals \( b_p \). Furthermore, a pair of port-2 and port-3 of the network model in Fig. 2 corresponds to the balanced output-port of the balanced SAW filter in Fig. 1, while port-1 corresponds to the unbalanced input-port. The network model shown in Fig. 2 can be also represented more simply by using a scattering matrix \( S = [s_{pq}] \) as follows.

\[
\begin{bmatrix}
  b_1 \\
  b_2 \\
  b_3
\end{bmatrix} =
\begin{bmatrix}
  s_{11} & s_{12} & s_{13} \\
  s_{21} & s_{22} & s_{23} \\
  s_{31} & s_{32} & s_{33}
\end{bmatrix}
\begin{bmatrix}
  a_1 \\
  a_2 \\
  a_3
\end{bmatrix}
\] (1)

2.3 Criteria of Balance Characteristics

In the balanced SAW filter, it is desirable that the output signals \( b_2 \) and \( b_3 \) from the balanced output-port, namely, a pair of port-2 and port-3 of the three-port network model in Fig. 2, have the same amplitude and 180 degrees phase difference through the pass-band. In order to evaluate those balance characteristics, we define two criteria that should be limited to small values[5]. The amplitude balance of the balanced SAW filter is evaluated with criterion \( E_1 \) in (2).

\[
E_1 = 20 \log_{10}(|s_{21}|) - 20 \log_{10}(|s_{31}|)
\] (2)

On the other hand, the phase balance of the balanced SAW filter is evaluated with criterion \( E_2 \) in (3).

\[
E_2 = \varphi(s_{21}) - \varphi(s_{31}) + 180
\] (3)

where, \( \varphi(s_{pq}) \) denotes the phase angle of \( s_{pq} \).

2.4 Criteria of Filter Characteristics

In order to evaluate the band-pass filter characteristics of the balanced SAW filter strictly, we have to segregate the differential mode signal from the common mode signal in the network model in Fig. 2. Therefore, according to the balanced network theory[8], the
differential mode signals \(a_d\) and \(b_d\) are derived from \(a_q\) \((q = 2, 3)\) and \(b_p\) \((p = 2, 3)\) as shown in (4). Similarly, the common mode signals \(a_c\) and \(b_c\) are also derived from them as shown in (5).

\[
\begin{align*}
    a_d &= \frac{1}{\sqrt{2}} (a_2 - a_3) \\
    b_d &= \frac{1}{\sqrt{2}} (b_2 - b_3) \\
    a_c &= \frac{1}{\sqrt{2}} (a_2 + a_3) \\
    b_c &= \frac{1}{\sqrt{2}} (b_2 + b_3)
\end{align*}
\]  

(4) (5)

From (4) and (5), the matrix \(S\) of conventional scattering parameters in (1) can be converted into the matrix \(S_{\text{mix}}\) of mix-mode ones as follows.

\[
S_{\text{mix}} = \mathbf{T} S \mathbf{T}^{-1} = \begin{bmatrix}
    s_{11} & s_{1d} & s_{1c} \\
    s_{d1} & s_{dd} & s_{dc} \\
    s_{c1} & s_{cd} & s_{cc}
\end{bmatrix}
\]  

(6)

where, matrix \(\mathbf{T}\) is given as follows.

\[
\mathbf{T} = \frac{1}{\sqrt{2}} \begin{bmatrix}
    \sqrt{2} & 0 & 0 \\
    0 & 1 & -1 \\
    0 & 1 & 1
\end{bmatrix}
\]

(7) (8)

By using the above mix-mode scattering parameters instead of conventional ones, we evaluate the band-pass filter characteristics of the balanced SAW filter in the same way with the unbalanced one[5]. Therefore, the standing wave ratios of the input-port \(E_3\) and the output-port \(E_4\) can be defined as (7) and (8). The attenuation \(E_5\) between the input-port and the output-port is also defined as shown in (9).

\[
\begin{align*}
    E_3 &= \frac{1 + |s_{11}|}{1 - |s_{11}|} \\
    E_4 &= \frac{1 + |s_{dd}|}{1 - |s_{dd}|} \\
    E_5 &= 20 \log_{10}(|s_{d1}|)
\end{align*}
\]  

(7) (8) (9)

\[E_3 = \frac{1 + |s_{11}|}{1 - |s_{11}|}, \quad E_4 = \frac{1 + |s_{dd}|}{1 - |s_{dd}|}, \quad E_5 = 20 \log_{10}(|s_{d1}|)\]

3 Problem Formulation

3.1 Design Parameters

In order to describe a suitable structure of the balanced SAW filter, we have to select appropriate design parameters such as the numbers of fingers for IDTs, the number of strips for SMSA, the width and the length of electrodes, and so on. Therefore, the design parameters of the balanced SAW filter usually take not only continuous values but also discrete values.

We represent the design parameters of the balanced SAW filter as \(\mathbf{x} = (x_1, \ldots, x_D)\). Besides, we specify the upper \(x_j^U\) and the lower \(x_j^L\) bounds for each of the design parameters \(x_j \in \mathbf{x}\) as follows.

\[
x_j^L \leq x_j \leq x_j^U, \quad j = 1, \ldots, D.
\]  

(10)

3.2 Objectives and Constraints

By using the above criteria \(E_h\) \((h = 1, \ldots, 5)\), we define three objective functions and six constraints.

First of all, the values of the above criteria \(E_h\) depend on both the frequency \(\omega\) and the design parameters \(\mathbf{x}\). Therefore, we choose a set of sample points \(\omega \in \Omega_P\) from the pass-band of the balanced SAW filter. Similarly, we choose two sets of sample points \(\omega \in \Omega_L\) and \(\omega \in \Omega_H\) respectively from the lower and the higher stop-bands of the balanced SAW filter.

Because the balanced SAW filter works as a band-pass filter, we define the following three objective functions \(f_m(\mathbf{x})\) \((m = 1, 2, 3)\) to be minimized by using the attenuation \(E_5 = E_5(\mathbf{x}, \omega)\) in (9).

\[
\begin{align*}
    f_1(\mathbf{x}) &= -\left( \sum_{\omega \in \Omega_P} E_5(\mathbf{x}, \omega) \right) \\
    f_2(\mathbf{x}) &= \sum_{\omega \in \Omega_L} E_5(\mathbf{x}, \omega) \\
    f_3(\mathbf{x}) &= \sum_{\omega \in \Omega_H} E_5(\mathbf{x}, \omega)
\end{align*}
\]  

(11) (12) (13)

We specify the upper \(U_h(\omega)\) and the lower \(L_h(\omega)\) bounds for the other criteria \(E_h(\mathbf{x}, \omega)\). Then the four of the six constraints \(g_k(\mathbf{x}) \leq 0\) \((k = 1, \ldots, 4)\) are given as shown in (14). The rest two constraints are given respectively as shown in (15) and (16).

\[
\begin{align*}
    g_k(\mathbf{x}) &= \sum_{\omega \in \Omega_P} \frac{E_k(\mathbf{x}, \omega) - U_k(\omega)}{|\Omega_P|} \leq 0 \\
    g_5(\mathbf{x}) &= \sum_{\omega \in \Omega_P} \frac{L_1(\omega) - E_1(\mathbf{x}, \omega)}{|\Omega_P|} \leq 0 \\
    g_6(\mathbf{x}) &= \sum_{\omega \in \Omega_P} \frac{L_2(\omega) - E_2(\mathbf{x}, \omega)}{|\Omega_P|} \leq 0
\end{align*}
\]  

(14) (15) (16)

3.3 Optimum Design Problem

From (10) ~ (16), we formulate the structural design of the balanced SAW filter as a constrained multi-objective optimization problem shown in (17).

\[
\begin{align*}
    \text{minimize} & \quad \{ f_1(\mathbf{x}), f_2(\mathbf{x}), f_3(\mathbf{x}) \} \\
    \text{subject to} & \quad g_k(\mathbf{x}) \leq 0, \quad k = 1, \ldots, 6. \\
    & \quad x_j^L \leq x_j \leq x_j^U, \quad j = 1, \ldots, D.
\end{align*}
\]  

(17)
4 Evolutionary Algorithm (GDE3)

4.1 Representation of Solution

GDE3 is usually used to solve the constrained multi-objective optimization problem over the $D$ ($D \geq 1$) dimensional real-parameters. GDE3 holds $N_F$ individuals, or the candidate solutions of the multi-objective optimization problem, in the population. As well as conventional real-coded GAs, every individual of GDE3 is coded as a $D$-dimensional real-parameter vector. Therefore, the $i$-th individual $x_i^G$ ($i = 1, \cdots, N_F$) included in the population of the generation $G$ ($G \geq 0$) is represented as follows.

$$x_i^G = (x_{i,1}^G, \cdots, x_{j,i}^G, \cdots, x_{D,i}^G) \quad (18)$$

where, $0 \leq x_{j,i}^G \leq 1$ ($j = 1, \cdots, D$).

As shown in Section 3, each design parameter $x_j \in \mathbb{X}$ of the balanced SAW filter takes either a continuous value or a discrete value. Therefore, in order to apply GDE3 to the structural design of the balanced SAW filter formulated as an optimization problem, we employ the following technique that converts an individual $x_i^G$ into the corresponding solution $x_i$[5].

In the regularized continuous search space of GDE3, each element of the individual $x_{j,i}^G \in x_i^G$ is restricted within the range between 0 and 1 as shown in (18). Therefore, each element $x_{j,i}^G \in x_i^G$ is converted into the corresponding design parameter $x_j \in \mathbb{X}$ when the values of the objective functions $f_m(x)$ and/or the constraints $g_l(x)$ are evaluated. If a design parameter $x_j \in \mathbb{X}$ takes a continuous value originally, the corresponding $x_{j,i}^G \in x_i^G$ is converted into the design parameter as shown in (19). On the other hand, if a design parameter $x_j \in \mathbb{X}$ takes a discrete value with an interval $e_j$, the corresponding $x_{j,i}^G \in x_i^G$ is converted into the design parameter as shown in (20).

$$x_j = (x_j^U - x_j^L) x_{j,i}^G + x_j^L \quad (19)$$

$$x_j = \text{round}\left(\frac{(x_j^U - x_j^L) x_{j,i}^G}{e_j}\right) e_j + x_j^L \quad (20)$$

where, round$(z)$ rounds $z \in \mathbb{R}$ to the nearest integer.

4.2 Procedure of GDE3

In the beginning of the procedure of GDE3, a set of individuals $x_i^G$ ($i = 1, \cdots, N_F$) are generated randomly as an initial population $x_i^G \in \mathbb{X}^G$ ($G = 0$).

Then, in each generation $G$ ($G = 0, \cdots, G_{\text{max}}$), GDE3 goes through each individual $x_i^G \in \mathbb{X}^G$, which is called the target vector, and generates trial vectors $u_i^G$ from $x_i^G$ with the genetic operator in (21). The genetic operator in (21) is equivalent to the strategy of DE called $\square$ DE/rand/1/bin $\square[6]$. Therefore, three individuals $x_{r1}^G$, $x_{r2}^G$ and $x_{r3}^G$ ($r_1 \neq r_2 \neq r_3 \neq i$) in (21) are selected randomly from the population $\mathbb{P}^G$.

$$\begin{align*}
    j_{\text{rand}} &= \text{rand}[1, D] \\
    \text{for}(j = 1; j \leq D; j = j + 1)\{ \\
    \text{if}(\text{rand}[0, 1] < C_R \lor j = j_{\text{rand}})\{ \\
    u_{j,i}^G &= x_{j,i}^G + SF (x_{j,r1}^G - x_{j,r2}^G) \\
    \} \text{else}\{ \\
    u_{j,i}^G &= x_{j,i}^G \}
    \}\}
\end{align*} \quad (21)$$

where, the subscript $j_r \in [1, D]$ is selected randomly. The scale factor $SF \in (0, 1]$ and the crossover rate $C_R \in [0, 1]$ are user-defined control parameters.

Each trial vector $u_i^G$ ($i = 1, \cdots, N_F$) is compared with the corresponding target vector $x_i^G$. Then, according to the following rules, either the trial vector $u_i^G$ or the target vector $x_i^G$ is selected as the member of the next population $\mathbb{P}^{G+1}$ for the time being[7].

- If both vectors are infeasible, then the trial vector $u_i^G$ is selected if it weakly dominates $x_i^G$ in the constraint violation space. Otherwise $x_i^G$ is selected.
- If one vector is feasible and the other is infeasible, then the feasible vector is selected.
- In the case that both vectors are feasible, then the trial vector $u_i^G$ is selected if it weakly dominates $x_i^G$ in the objective function space. On the other hand, if $x_i^G$ dominates $u_i^G$, then $x_i^G$ is selected. If neither vector dominates each other in the objective function space, then both vectors are selected.

After the above selection in each generation, the size of the next population $\mathbb{P}^{G+1}$ may have increased over the original size $N_F$. If that is the case, the size of the population $\mathbb{P}^{G+1}$ is decreased back to the original size based on a similar selection approach used in NSGA-II[9]. Exactly speaking, the individuals of the population $\mathbb{P}^{G+1}$ are sorted based on non-dominance and crowdedness. Then the inferior individuals according to these measurements are removed from $\mathbb{P}^{G+1}$ to decrease the size of $\mathbb{P}^{G+1}$ to $N_F$.

5 Computational Experiment

5.1 Setup of Experiment

As an instance of the multi-objective optimum design problem, a suitable structure of the balanced SAW filter illustrated in Fig. 1 is considered. In order to describe the structure of the balanced SAW filter, we have selected $D = 11$ design parameters $x_j \in \mathbb{X}$ ($j = 1, \cdots, D$) as shown in Table 1. Besides the
Table 1: Design parameters of balanced SAW filter

<table>
<thead>
<tr>
<th>$x_j$</th>
<th>$x_j^L, x_j^U$</th>
<th>$e_j$</th>
<th>design parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>[200, 400]</td>
<td>–</td>
<td>overlap between electrodes</td>
</tr>
<tr>
<td>$x_2$</td>
<td>[10.0, 40.0]</td>
<td>0.5</td>
<td>number of fingers of IDT-R</td>
</tr>
<tr>
<td>$x_3$</td>
<td>[10.5, 40.5]</td>
<td>1.0</td>
<td>ditto of IDT-T</td>
</tr>
<tr>
<td>$x_4$</td>
<td>[1.0, 4.0]</td>
<td>1.0</td>
<td>ditto of modulated IDT</td>
</tr>
<tr>
<td>$x_5$</td>
<td>[50.0, 300.0]</td>
<td>10.0</td>
<td>number of strips of SMSA</td>
</tr>
<tr>
<td>$x_6$</td>
<td>[0.2, 0.8]</td>
<td>–</td>
<td>metallization ratio of IDT</td>
</tr>
<tr>
<td>$x_7$</td>
<td>[0.2, 0.8]</td>
<td>–</td>
<td>ditto of SMSA</td>
</tr>
<tr>
<td>$x_8$</td>
<td>[1.0, 1.1]</td>
<td>–</td>
<td>pitch ratio of SMSA</td>
</tr>
<tr>
<td>$x_9$</td>
<td>[0.9, 1.0]</td>
<td>–</td>
<td>ditto of modulated IDT</td>
</tr>
<tr>
<td>$x_{10}$</td>
<td>[1.9, 2.1]</td>
<td>–</td>
<td>finger pitch of IDT</td>
</tr>
<tr>
<td>$x_{11}$</td>
<td>[3900, 4000]</td>
<td>–</td>
<td>thickness of electrode</td>
</tr>
</tbody>
</table>

The design parameters $x_j \in \mathbb{R}$, Table 1 shows their upper $x_j^U$ and lower $x_j^L$ bounds. Furthermore, if a design parameter $x_j \in \mathbb{R}$ has to take a discrete value, the interval $e_j \in \mathbb{R}$ is also described in Table 1.

The objective functions $f_m(x)$ and the constraints $g_k(x)$ shown in (17) are evaluated at 401 sample points $\omega \in \Omega_F \cup \Omega_P \cup \Omega_R$ within the range between 850[MHz] and 1080[MHz]. The pass-band is also selected to the range from 950[MHz] to 980[MHz].

As the stopping condition of GDE3, the maximum generation is limited to $G_{\text{max}} = 800$. The control parameters of GDE3 are given as follows: the population size $N_P = 200$, the scale factor $S_F = 0.9$ and the crossover rate $C_R = 0.9$. These values are decided considering the result of the empirical study about the control parameters of GDE3[10].

Incidentally, the program of GDE3 including the simulator for the balanced SAW filter shown in Fig. 1 is coded by MATLAB. The program spends about one hour for one run on a PC (CPU: Intel(R) Core 2).

5.2 Result of Experiment

Figure 3 depicts all individuals of the populations of different generations in the objective function space. In Fig. 3, infeasible individuals are denoted by cross symbol (×), while feasible ones are denoted by circle symbol (○). Furthermore, non-dominated feasible individuals are denoted by blue circle, while dominated feasible ones are denoted by red circle.

Comparing the populations shown in Fig 3, we can observe the great progress of the multi-objective search has been made by GDE3. First of all, all individuals of the initial population ($G = 0$) are worse and infeasible. However, the objective function values of individuals are improved and a lot of feasible individuals are found after $G = 200$ generations. Finally, at the generation $G_{\text{max}} = 800$, all individuals become feasible and they are non-dominated each other.

5.3 Analysis and Discussion

Recently, Principal Component Analysis (PCA) has been used successfully to assess a set of the Pareto-optimal solutions obtained by EMO methods[11]. In order to clarify the tradeoff relationship among the three objective functions of the above optimum design problem, we have also applied PCA to the set of the non-dominated solutions obtained by GDE3.

Table 2 shows the result of PCA in which eigenvectors $\alpha_m$ ($m = 1, 2, 3$) and accumulated proportions $\alpha_m$ are listed. Because $\alpha_2 > 90\%$ holds in Table 2, we may pay attention only to the first and the second principal components. Furthermore, considering the objective functions corresponding to the most positive and the most negative element of the first principal component $\zeta_1$, we can see that $f_1(x)$ and $f_3(x)$ are the two most critically conflicting objectives, whereas $f_2(x)$ seems to be redundant for the three-objective optimization problem in (17).

6 Conclusion

A multi-objective optimum design method for balanced SAW filters was proposed. First of all, the structural design of the balanced SAW filter was formulated as a constrained three-objective optimization problem. In order to obtain various Pareto-optimal solutions, a recent EMO called GED3 was applied to the optimization problem. Furthermore, in order to assess the set of the non-dominated solutions obtained by GDE3, PCA was employed. As a result, we could find that the two of the three objective functions were clearly conflicting. That might be useful information for the engineers of the balanced SAW filter.

Future work will focus on the further investigation of the set of the Pareto-optimal solutions in the design parameter space. Thereby we would like to clarify the relationship between the structure and the performance of the balanced SAW filter.

Acknowledgements: The research was supported in part by the Grant-in-Aid for Scientific Research (C) (Project No. 21560432) from Japan Society for the Promotion of Science (JSPS).
Figure 3: Population of generation $G$ ($N_P = 200$)

References:


