Global Optimization Method to Locate Multiple Local Optima with Response Surface Approximation Methodology*

Qinzhong SHI**, Ichiro HAGIWARA**
and Futoshi TAKASHIMA***

A considerable number of functional evaluations may be required in the process of optimization. Although approximation models constructed by response surface methodology can significantly reduce functional evaluations, the design accuracy may be strongly dependent on the type of activation functions and designs used. In this paper, we propose techniques to search the design space containing the global optimal design using designs by conditioned random seeds, and techniques for determining more accurate approximations by employing a sequential approach called the most probable optimal design (MPOD) method. The MPOD method is a response surface methodology based on the holographic neural network, which uses the exponential function as an activation function. In the MPOD method, extrapolation is employed to make the technique available for general application in structural optimization. The formula to estimate the necessary functional evaluations under certain conditions is expressed. Application examples show that the MPOD method is an effective methodology, which is expected to determine the optimal design with multiple local optima.

**Key Words:** Nonlinear Problem, Neural Network, Finite-Element Method, Numerical Analysis, Buckling, Optimal Design, Response Surface Methodology, Extrapolation Technique, Global Optimal Design

1. Introduction

Due to the recent developments in computer science and software technology, fluid analysis, nonlinear structural analysis and their combined analysis has become possible for large-scale models and will eventually be employed in manufacturing design(1). The importance of being able to make rational and correct decisions during the early stages of design is currently being emphasized. However, crashworthiness analysis requires considerable time for one functional evaluation even using supercomputers. It is vital that rapid evaluation models of performance be used instead of computationally time-consuming simulation tools and key techniques in optimal design. For a complex optimal design problem, teams of design engineers must achieve maximum product performances to implement the concurrent design. Optimization techniques which cover several design disciplines (multiple discipline optimization-MDO) have been hot topics recently(2). Sensitivity-based optimization approaches, which sequentially construct local approximations, are widely employed and provide valuable contributions to linear optimization problems(3). Unfortunately, sensitivity approaches are not available for nonlinear optimization problems. For this purpose, response surface methodology (RSM) works effectively to achieve these design optimizations. Instead of the time-consuming finite-element method (FEM) analysis, robust approximation models can be constructed to evaluate responses rapidly in the design space of interest by RSM, when the response surface of the structure is infiltrated with computational noise. Research has proven that RSM is a valuable tool for optimization problems in which the designer must work with functions with non-
smoothing response surfaces. RSM can also be used to construct global and midrange approximations for functions in structural optimization to achieve rapid convergence. The other advantage of RSM is that the designer is able to obtain an overall insight into the behavior of complex phenomena over a wide range of design space. White used RSM in their study of passenger-car crushworthiness and Giunta applied RSM to smooth the noisy response surface in high-speed civil transportation design. However, several key issues must be dealt with to develop the response surface modeling. One issue for RSM may be that the exact form of activation response surface is not available for general applications and the knowledge of design performance necessary to select the activation of RSM. Most response surfaces are expressed in terms of polynomials. Unfortunately, this is not always accepted as a proper activation function for response surfaces as a real response surface and may not provide a satisfactory solution, particularly for wide ranges of design space with heavy nonlinearity. This limits the validity of RSM to only cover a part of the design space. Another problem in using RSM is the difficulty in predicting the error incurred using response surface function values rather than the “true” values from the analysis code. Therefore, the extensive RSM research is focused on the selection of designs in design space, such as central composite designs (CCD), and D-optimality.

A neural network is considered as the most feasible activation response function to RSM. Carpenter et al. compared the multilayer feed-forward neural networks with polynomials to demonstrate that neural networks show no appreciable difference in the accuracy of the two models for special problems. Hagiwara et al. initially employed the holographic neural network for the activation function of the response surface and compared it to the multilayer feed-forward neural network to demonstrate that the former type of neural networks had a much higher approximation accuracy and training speed. The training of a holographic neural network costs about 1/1000 the computational time compared to that of a multilayer feed-forward neural network. The application of the holographic neural network (HNN) to search all the local optimal designs within the feasible design space, and techniques for determining more accurate approximations have been investigated, on the assumption that the designer approximately knows the location of the local optimal. The minimum required number of initial designs for training the HNN and adding designs concerning extensive training to improve the approximation accuracy near the local optimal have been proposed.

In this paper, a sequential approach using the HNN is introduced for the activation of the response surface to determine the global optima efficiently, and a mapping method of extrapolation is proposed to make the technique available for general application in structural optimization. Since the exact form of response surface remains unclear in general application, and the accuracy of the function approximation near the global optimal design is very important, techniques to search the design space containing the global optimal design are examined. For the purpose of determining the global optimal design, techniques for searching the most likely global optimal design over the entire design space, and techniques for determining a more accurate approximation near the global optimal designs using the extrapolation ability of the HNN are proposed. In this paper, we present the proposed approach, termed the most probable optimal design (MPOD) method to determine one local optimal design, which has the highest probability in the design space. The MPOD is a sequential procedure composed of two steps. The first step is employed to determine the most likely global optimal design over the entire design space by conditioned random seeds, which control the mutual distances between every two designs, to scan the wide design space rapidly. The second step is to verify and improve the accuracy of the approximation of the objective function and constraints at optimal design. The trade-off between the approximation accuracy and computational cost can be well balanced by the preset threshold of the mutual distance of designs. A design example of crashworthiness for vehicular structure demonstrates the validity and utility of this method.

2. Response Surface Methodology

2.1 Theoretical background of response surface methodology

RSM comprises a group of statistical techniques for empirical model construction and model exploitation by careful design selection and corresponding analysis or experiments to seek the function relation of response/design performance to design variables. Using RSM, it is possible to construct a trade-off approximation model with significantly fewer functional evaluations in order to determine the global optimal design in the feasible design space. The true response surface is given as

$$\tilde{y} = \eta(x),$$ (1)

where \(\eta(x)\) is the true response and \(x\) is the design variable vector. The true response is generally unknown and may be presumed to be an approximated function \(f(x)\) in the design space \(D(x)\),
\[ y(x) = f(x) + \delta(x), \quad (2) \]

where \( \delta(x) \) is the systematic (bias) error, which is the bias between the approximated function and the true functional response.

Considering the calculation or experimental random error \( \varepsilon \), the functional relationship at each experimental design \( x_i \), may be precisely expressed by Eq. (3).

\[ y(x_i) = \eta(x_i) + \varepsilon_i = f(x_i) + \delta_i + \varepsilon_i \quad (i = 1, 2, \cdots, p), \quad (3) \]

where \( p \) is the number of experimental designs and \( \theta \) is the model parameter. In Eq. (3), the bias error of the approximation model \( \delta \) may result in a fatal weakness of the approximation accuracy which is strongly dependent on the activation response function used for RSM. \( \varepsilon \) is always assumed to be a zero mean value, statistically independent, with constant variance \( \sigma^2 \) and is normal distributed, that is, \( \varepsilon \sim N(0, \sigma^2) \). A great simplification in the estimation of the response function is the linear combination of activation functions, that is,

\[ f(x) = \sum_{k=1}^{m} a_k \phi_k(x), \quad (4) \]

where \( \phi_k(x), \quad (k = 1, 2, \cdots, m) \) are the activation functions, \( \alpha_k (k = 1, 2, \cdots, m) \) are the model parameters, and \( m \) is the number of activation functions. The estimator of model parameters \( \hat{\alpha}_k (k = 1, 2, \cdots, m) \) can be calculated by the maximum likelihood estimator from Eq. (5) as

\[
P(\{A\} | Y) \propto \exp \left( -\frac{1}{2\sigma^2} \left( Y - [\{W\}][\{A\}] \right)^T \left( [W]^T [W] \right)^{-1} \left( Y - [\{W\}][\{A\}] \right) \right)
= \max \quad (5)
\]

The solution of Eq. (5) is equivalent to the solution of Eq. (6),

\[ f(\{\hat{A}\}) = (Y - [\{W\}][\{A\}])([W]^T [W])^{-1} [W]^T [Y] \Rightarrow \min, \quad (6) \]

where \( t \) is the transpose, \( W \) is the covariance matrix of error \( \varepsilon \) and the parameter vector is

\[ \{A\} = [a_1, a_2, \cdots, a_m]. \quad (7) \]

Let

\[ \{\phi\} = [\phi_1(x_1), \phi_2(x_1), \cdots, \phi_n(x_1); \phi_1(x_2), \phi_2(x_2), \cdots, \phi_n(x_2); \cdots; \phi_1(x_p), \phi_2(x_p), \cdots, \phi_n(x_p)]. \quad (8) \]

The maximum likelihood solution of \( \{A\} \) is

\[ \{\hat{A}\} = ([\{\phi\}]^T [W]^T [W])^{-1} [W]^T [Y]. \quad (9) \]

The distribution of the solution \( \hat{\alpha}_k (k = 1, 2, \cdots, m) \) is also normally distributed with the deviation \( \bar{\varepsilon} \)

\[ \{\hat{A}\} \sim N(\{A\}), ([\{\phi\}]^T [W]^T [W])^{-1}. \quad (10) \]

We can see from Eq. (10), that the estimator of model parameters is statistically efficient. Here, we must note that the bias error is completely ignored. The deviation of the estimator is proportional to the experimental error \( \varepsilon_k (k = 1, 2, \cdots, p) \). The proportional coefficient \( ([\{\phi\}]^T [W]^T [W])^{-1} \) is dependent on the activation function and the designs \( x_k \). The careful selection of designs to make the proportional coefficient minimum and independent between each column in Eq. (8) is the key point of the response surface. The previously described methods, CCD and D-optimality, are based on this consideration. CCD is used to select fully orthogonal designs. D-optimality is used to determine that the proportional coefficient is minimum. D-optimality is considered to be the most flexible criteria for design selection and is suitable for any feasible design space \( D(x) \).

2.2 Holographic neural network for response surface to reduce bias error

A holographic neural network adopted as the activation function for the response surface has been reported in previous work of the authors\(^{19}\). Here, the theory of the holographic neural network for the response surface is briefly introduced.

In general, the form of stimulus of real data \( x = [\xi_1, \xi_2, \xi_3, \cdots, \xi_n] \) may be expressed by transforming the scalar values to complex space by the following nonlinear transformation.

\[ \phi_k(\xi_k) = -\lambda^k e^{i\theta^k}, \quad (k = 1, 2, 3, \cdots, n) \quad (11) \]

\[ \theta^k = \frac{2\pi}{1 + e^{-c_k(\xi_k-a_k)}} \quad (k = 1, 2, 3, \cdots, n), \quad (12) \]

where \( \mu_k \) and \( \sigma_k \) are the mean value and standard deviation of \( \xi_k \), respectively, and \( \lambda^k \) defines a vector magnitude bound within the unit circle (0 to 1), expressing a weighting or dominance for each element of the stimulus field \( (k = 1, 2, 3, \cdots, n) \). If no design variable dominates, \( \lambda^k = 1 \) \( (k = 1, 2, 3, \cdots, n) \). \( c_k \) is the coefficient to adjust the transformation slope shown in Fig. 1.

A similar transformation can be performed within the response field \( y \), where the scalar data represents information expressed within a real number domain,

\[ r = f \tau e^{i\theta} = fr(y) \quad (13) \]

where \( fr \) is transformation functions, which are always nonlinear functions. If \( p \) designs are trained, then \( \phi_k(\xi_k) \) and \( r \) are summarized as the following

\[
\begin{align*}
\text{Fig. 1 Transform function of phase angle}
\end{align*}
\]
matrix formulations.

\[
[J_\psi] = \begin{bmatrix}
\lambda_1 e^{i\theta_1} & \lambda_2 e^{i\theta_2} & \cdots & \lambda_n e^{i\theta_n} \\
\lambda_1^* e^{-i\theta_1} & \lambda_2^* e^{-i\theta_2} & \cdots & \lambda_n^* e^{-i\theta_n} \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_1^* e^{-i\theta_1} & \lambda_2^* e^{-i\theta_2} & \cdots & \lambda_n^* e^{-i\theta_n}
\end{bmatrix}
\]  

(14)

\[
Y = \{Y^* e^{i\phi_1}, Y^* e^{i\phi_2}, \ldots, Y^* e^{i\phi_n}\}
\]  

(15)

where \( \theta_k \), \( \phi_k \) and \( \lambda_k \) represent the mapped phases and magnitudes of the stimulus and response, respectively. \( k \) is the element index and \( j \) is the index of trained designs index.

The training operation involves mapping the stimulus \( \{\phi\} \) and response \( \{Y\} \). These mappings can be expressed by linear functions in the parameters of Eq. (4).

\[
r = \sum_{k=1}^{n} a_k \phi_k(\xi),
\]

(16)

where \( \phi_k(\xi) \) is the activation function decided by Eq. (11) and \( a_k \) are the complex parameters of the model. The maximum likelihood estimator of model parameters can be calculated from Eq. (9). Let the covariance matrix of experimental error be a unit matrix \( \{W\} = \{I\} \), then the estimators of parameters become

\[
\{\hat{A}\} = \left(\{\phi\}^H \{\phi\}\right)^{-1}\{\phi\}^H \{Y\},
\]

(17)

where \( H \) stands for the Hermitian transpose (conjugate transpose).

To solve parameter \( \{\hat{A}\} \) in Eq. (17), it is necessary to calculate the inverse matrix. When the number of activation functions is large, it will lead to considerable calculation time and may result in calculation error. The following iteration process is used to save calculation time.

\[
\{\hat{A}\} = \{\{\phi\}^H \{Y\}\}
\]

(18)

The enhancement of parameters will be performed by

\[
\{\Delta \hat{A}\} = \{\{\phi\}^H \{Y\}\} - \{\{\phi\} \{\hat{A}\}\}
\]

(19)

\[
\{\hat{A}\}_{\text{new}} = \{\hat{A}\} + \{\Delta \hat{A}\}
\]

(20)

The iteration may be terminated when the error between the approximating response and the true response is orthogonal to every column vector of \( \{\phi\} \) in a complex domain. In practice, it is not essential that the iteration is terminated when the error becomes smaller than the preset threshold value, \( \epsilon_{ps} \).

\[
J((\hat{A}_{\text{new}})) = \{(Y) - \{\phi\} (\hat{A}_{\text{new}})\}^H ((Y) - \{\phi\} (\hat{A}_{\text{new}})) \leq \epsilon_{ps}
\]

(21)

where \( \epsilon_{ps} \) is the threshold value to terminate the iteration. It is clear from the algorithm that the holographic neural network will be entirely convergent, independent of the initial value. The error vector \( \{v_{nu}\} \) can be calculated by Eq. (22).

\[
\{v_{nu}\} = \{(Y) - \{\phi\} (\hat{A})\}
\]

\[
= \{(Y) - \{\phi\} (\{\phi\}^H \{\phi\})^{-1} \{\phi\}^H \{Y\}\}
\]

(22)

The error vector \( \{v_{nu}\} \) includes the bias error and experimental error, as stated previously. The statistical distribution can be determined from Eq. (10) if the experimental error is independent and in the normal distribution with the zero mean value.

\[
\{\hat{A}\} \sim N(\{A\}, \{\phi\}^H \{\phi\}^{-1} \{\phi\}^H)
\]

(23)

The predicted response can easily be calculated from Eq. (16).

\[
\{r\} = \{\phi(x_d)\}^H \{\hat{A}\}
\]

(24)

Generally, the bias error is based on the number of orders \( n \) of the activation function. The greater the number of orders of the activation function, the smaller the bias error becomes. Conventional artificial neural network models are reasonably restricted in terms of the number of associations that may be accurately encoded into a neural network. Within the holographic neural network, limitations of mapping accuracy can be largely overcome through a preprocessing operation involving the generation of higher-order product terms. For example, for two design variables \( n = 2 \), \( s = \{\lambda e^{i\theta_1}, \lambda e^{i\theta_2}\} \) can be increased to 12 terms \( n = 12 \) if performed to the second order.

\[
S = \{\lambda e^{i\theta_1}, \lambda e^{i\theta_2}, \lambda e^{i2\theta_1}, \lambda e^{i2\theta_2}, \lambda e^{-i\theta_1}, \lambda e^{-i\theta_2}, \lambda e^{-i2\theta_1}, \lambda e^{-i2\theta_2}, \lambda \lambda e^{i(\theta_1 + \theta_2)}, \lambda \lambda e^{i(\theta_1 - \theta_2)}, \lambda \lambda e^{-i(\theta_1 + \theta_2)}, \lambda \lambda e^{-i(\theta_1 - \theta_2)}\}
\]

(25)

Furthermore, the extrapolation of the approximated response function may be easily performed with the trained holographic neural network by presetting the coefficient \( c_k \) in Eq. (12) to a small value to adjust the slope to be smaller. Figure 1 shows the effect of the coefficient. In Fig. 1, the experimental designs used for training the neural network are located within the range \([0, 1]\), the dotted line expresses the transform when \( c_k = 1.0 \), the solid line expresses the transform when \( c_k = 0.5 \). We can observe from the figure that if the predicted designs are outside the trained design space, then the predicted function values outside the trained design space may be considered to be the same values as those at trained design edges 0 or 1 for the case of \( c_k = 1.0 \). In contrast, the predicted function values outside the trained design space may keep the trend beyond the trained design edges 0 or 1 for the case of \( c_k = 0.5 \). Figure 2 shows the effect of extrapolation. In this figure, space “1” expresses the design space inside which designs are used for approximation by training, and space “D” expresses the entire feasible design space. Approximation using designs in space “1” may extrapolate the values of space “D”. This is an important issue for reducing functional evaluations and some advantages of this are: 1) design spaces with irregular shapes can be considered, 2) any number of experimental designs can be considered, and 3) searching for reasonable optimal designs over a wide space becomes possible.
3. Optimal Design for Problems with Multiple Local Optima-Most Probable Optimal Design Method (MPOD)

3.1 Procedure for MPOD

The following procedure is proposed to determine the most robust global optimal design with a small number of functional evaluations over a wide design space. These considerations arise from the concept that concentrated designs similar to the optimal-like design and scattered designs very unlike the optimal-like design are used for approximating the desired function. A two-step sequential optimization and approximation procedure is proposed to determine the optimal design. The first step is to scan the entire design space to determine where the global optimal design exists, most probably by designs selected using conditioned random seeds. The second step is to increase the accuracy of the most-likely global optimal design by inputting additional information where required. The details regarding these two steps are described below. The optimal problem is defined as

\[ \text{max } f(x), \]  

subject to: \( g_i(x) \leq 0 \), \( i = 1, 2, \ldots, m \),

where \( m \) is the number of constraints.

**Step 1:** To search the most-likely global optimal design over the entire design space.

Selecting initial trial designs \( x_k \), \( k = 1, 2, \ldots, N \), where \( N \) is the initial number of design points. In the case where \( N = (n+1)(n+2)/2 \), for example, \( n \) is the number of design variables. However, one should understand that this is not the only choice. Actually, the determination of the \( N \) value depends on the scale of the problem, whether the problem has a large number of design variables and the computation time of each functional evaluation. For instance, the formulation of center composite design points \( N = 2n + 1 \) is recommended for problems with a large number of design variables. It should be emphasized that the designs produced by conditioned random seeds are employed to search the entire design space. The uniformly distributed conditioned random seeds of designs are used to provide a good balance of search in the design space. Design conditions are required to satisfy Eq. (28).

\[
d_n = \min \left( \frac{\| x_i - x_j \|}{\| x_{cc} \|} \right),
\]

\( 1 \leq i, j \leq N, i \neq j, m = 1, 2, \ldots, p \) \hspace{1cm} (28)

where \( x_{cc} \) is the normalization parameter which can be selected to be the geometrical center of design \( x_1 \).

**Step 2:** To verify and improve the accuracy of optimal design and its performance. The accuracy check is based on

\[ d = \frac{\| x_{op}(k) - x_{op}(k-1) \|}{\| x_{cc} \|} \leq \varepsilon \]

Thus far, the above sequential approximation was used to determine the approximation global optimal design location.
one-dimensional example is employed to express the estimator. Assume the size of design space $\Omega$ is $d_0$, and the size of the entire feasible design space is $D$. We define the space $\Omega$ as an optimal cluster, which is mathematically defined by Eq. (31).

$$d_0 = \min \left\{ \| x_e - x_c \| : \| x_c \| \leq \| x_e \| \right\}$$  (31)

$$f(x_e) > f(x_c) ; x_c \in \Omega ; x_e \not\in \Omega$$  (32)

which indicates that the function values within the optimal cluster $\Omega$ are larger than those outside the optimal cluster. $x_e$ is the edge point of the optimal cluster, $x_c$ is the geometrical center of optimal cluster, and $d_0$ expresses the size of the optimal cluster. If the size of the optimal cluster is known, it is estimated that the maximum number of designs for determining the optimal cluster with probability unity is decided by Eq. (33).

$$P = \int \left( \max \left\{ \frac{2 \| x_e - x_c \|}{d_0} \right\} \right)$$  (33)

where $x_e$ expresses the group points on the edge of design space, $x_c$ is the geometrical center of the design space, and $\int$ expresses the rounded integer. In this case, if $\varepsilon_1 = d_0$ in Step 1, it can be moved to Step 2 with less than $P$ designs.

4. Numerical Examples

4.1 Comparison of approximation by holographic neural network and polynomials

The comparison of the holographic neural network with polynomials for the function approximation is performed by the following multiple peaks function, which is shown in Fig. 4.

$$f(x) = 10^x x^8 (x^2 - 2)^8 e^{-x^2}$$

subject to : $-4.0 \leq x \leq 3.0$

For the above numerical example, the number of training designs is 15 and they are evenly spaced over the design space. The approximation accuracy of the holographic neural network and polynomials is shown in Fig. 5. In the figure, the horizontal axis expresses the orders of polynomials (a) or orders of holographic
neural network (b), the vertical axis expresses the standard variance calculated with the 200 designs evenly spaced in the design space \([-4.0, 3.0]\).

We can see from the results that the holographic neural network is much more accurate than the polynomials and maintains good accuracy after the ninth order of expansion. The standard deviation of the holographic neural network at the ninth order is 0.3, while those of polynomials at the 14th order is higher than 1.0. In other words, a larger number of coefficients must be selected for the polynomials by Eq.(10), and thus, more experimental designs must be used compared to that of the holographic neural network. Therefore, the holographic neural network is a more suitable activation response function for this problem than polynomials.

4.2 Numerical example with holographic neural network extrapolation application

To demonstrate the efficiency of this method, a numerical example of multiple local maxima with constraints is used to search the global optimal design in the feasible design space. The optimization problem is defined as

\[
\begin{align*}
\text{max } & f(x_1, x_2) \\
\text{subject to } & (-2, -2) \leq (x_1, x_2) \leq (2, 2), \\
& g_1(x) = \frac{15}{80} x_1^3 - \frac{1}{8} x_1 + x_2 - 2 \leq 0 \\
& g_2(x) = \frac{1}{2} x_1 - x_2 - 1 \leq 0
\end{align*}
\]

where

\[
f(x_1, x_2) = 3(1-x_1)^2e^{-x_1^2-(x_2+1)^2} \\
10\left(\frac{x_1}{5} - x_2^2 - 1\right) - e^{-(x_1^2 + x_2^2)} - \frac{1}{3} e^{-(x_1+1)^2-(x_2+1)^2}
\]

Three local optimal designs, marked 1, 2, 3, exist within the feasible design space depicted in Fig. 6. It is a contour plot, and the feasible design space is enclosed within the shadowed area. The following two cases are performed to investigate the performances of convergent accuracy and the number of functional evaluations.

(1) Case of large deviation (\(\sigma = 5\Delta d\)): A sequential approximation is performed using the proposed two steps stated previously. Value of \(\beta\) is set to 0.6. The method of training the holographic neural network is: to start with expansion terms of 2, and training accuracy of 0.02, which if not satisfied after 1000 epochs, is gradually increased to a higher-order expansion by two orders. The thresholds of Eqs.(29) and (30) are \(\epsilon_1 = 0.15\) and \(\epsilon_2 = 0.01\). Step 1 is terminated after nine iterations (corresponding to 22 functional evaluations) of random design searches. Then, Eq.(30) is satisfied after two iterations and the total number of functional evaluations until convergence is 24. The approximation optimal design is \(x_{opt}(k) = (-0.013, 1.583)\), \(f(x_{opt}(k)) = 8.106\). The history of each iteration is illustrated in Fig. 7, in which the horizontal axis stands for the number of functional evaluations and the vertical axis stands for the distance (\(d_k\)) between the original position (0, 0) and the approximation optimal design.

(2) Case of small deviation (\(\sigma = 2\Delta d\)): The training procedure and stop thresholds are similar to the case of large deviation. Step 1 is terminated after six iterations (corresponding to 16 functional evaluations) of random design searches. Then, Eq.(30) is
satisfied after two iterations and the total number of functional evaluations until convergence is 18. The approximation optimal design is $x_{opt}(k)=(-0.013, 1.584)$, $\tilde{f}(x_{opt}(k))=8.105$. The history of each iteration is illustrated in Fig. 8, in which the horizontal and vertical axes have the same meaning as those of Fig. 7.

Further inspection of the global approximation between the above two cases is performed using 1600 (40×40) lattice design points covering the entire design space. The deviations of cases (1) and (2) are 1.12 and 1.32, respectively. That is, a larger deviation of Gauss random designs gives slower convergence but higher accuracy of approximation over the entire design space. On the other hand, a smaller deviation of Gauss random designs gives rapid convergence but lower accuracy of approximation over the entire design space.

Let us confirm the effect of Eqs. (33) and (34). If the threshold termination in Step 1 is set to $\varepsilon_1=0.15$, the number of designs is 12 and 22 in cases $\alpha=2$ and $\alpha=5$, respectively. If the threshold termination in Step 1 is set sufficiently small, say, $\varepsilon_1=1.0 \times 10^{-3}$, then the history of convergence is depicted in Fig. 9 and the number of functional evaluations in Step 1 becomes 36. However, on the basis of the result shown in Fig. 9, we know that there is no further improvement in determining the global optimal design after 36 functional evaluations, except to increase the probability. On the other hand, $d_{th}=0.9$ can be determined for this problem, and Step 1 is terminated when $\varepsilon_1=0.9$ at eight designs. In practice, however, this is not known; therefore, if one attempts to search for the global optimal design including small clusters, one is expected to select a stricter termination threshold in Step 1 and a larger deviation of Gauss random designs.

4.3 Vehicle side member component optimal design

The vehicle side member (component parallel to the central axis of vehicle) plays a role in energy absorption while the crashworthiness and energy absorption of the vehicle is determined by its size, shape and welding. In this example, the square cross-area, a perfectly straight side member with uniform thickness is investigated and reinforcement of the component is considered as a way to increase the energy absorption. The size and material property of the member is illustrated in Fig. 10(a), and the form of reinforcement is depicted in Fig. 10(b). All degrees of freedom at the bottom end are rigidly fixed, and at the top end, a rigid mass of 500 kg, and velocity of 54 km/h are used as a load to simulate a crash. The load-displacement behavior of the member while crashing is calculated by the finite-element method (FEM) solver LS-DYNA3D [13], in which 2700 shell elements and 2754 nodes are used. The design variables include the thickness of the baseplate ($t_b$), the upperplate ($t_u$) and the location of reinforcement ($z$). The total weight of the component is constant ($w=780$ g), therefore only two design variables are independent. The objective function is the energy absorption of the component, 10 ms after crashing.
The mathematical definition of the problem is
\[
\begin{align*}
\text{max } & f(t_1, t_2, z) \\
\text{subject to: } & 0.5 \leq t_1, t_2 \leq 1.0, 1.5 \leq z \leq 250 \\
& t_2 = 1.5 + (1 - t_1) \\
& 2.0
\end{align*}
\]

The training procedure and stop thresholds are identical to the case of large deviation, that is, \( a = 5, \beta = 0.6, \) and \( \epsilon_1 = 0.15, \epsilon_2 = 0.01. \) Step 1 is terminated after four iterations (corresponding to 12 functional evaluations) of random seed searches. Then, Eq. (30) is satisfied after two iterations and the total number of functional evaluations until convergence, is 14. The approximation optimal is \((1.21, 1.39, 160), f(x_{opt}(k)) = 8484.11 \text{ kJ}, \) which is 11.6% higher than that of the original design \((1.00, 1.50, 0.0). \) The contour plot of the approximation function is depicted in Fig. 11. In the figure, the open circle indicates the optimal design using the proposed MPOD method and the plus sign expresses the designs used for approximation. The comparison between the approximation function and FEM values with \( t_1 = 1.21 \) and \( t_2 = 1.39 \) are fixed while the changing reinforcement location \( z \) is plotted in Fig. 12. It can be seen from the figure, and also mentioned in Ref. (14), that the FEM value varies, because of the heavy nonlinearity of the crashworthiness problem (14). A smooth and robust approximation function is obtained by the holographic neural network approximation and an optimal design is realized successfully (16). The MPOD approach based on the holographic neural network has a robust property against calculation noise. However, the deviation of the coefficient is dependent on the design distribution. This topic will be discussed further in future works.

5. Conclusion

In this paper, we presented a novel method to solve the global optimal design within a feasible design space using response surface methodology based on holographic neural network activation. The sequential approximation procedure and extrapolation of the holographic neural network were proposed for irregular design space and arbitrary designs. The detailed summaries of the issues raised in this paper are as follows.

1) Response surface methodology (RSM) is employed to construct the approximation functions which normally contain systematic and random errors. The state-of-the-art research focuses on composing a group of statistical techniques for empirical model building and model exploitation by careful design selection, such as orthogonal design, central composite design, and D-optimality design to minimize the confident region of the estimated coefficient. A new type of flexible activation function similar to the neural network is expected as a potential candidate to reduce systematic errors, which are much more significant in RSM.

2) In this paper we deal with the systematic errors and a numerical example demonstrates that the holographic neural network is more feasible to approximate the multiple peaks function than the polynomials.

3) The extrapolation can be performed by adjusting the transformation coefficient which expresses the slope of the transformation function of the holographic neural network. It is shown that higher orders of the holographic neural network can be realized by the products of fundamental terms to improve the approximation capability.

4) It is illustrated that the proposed sequential approximation procedure involves two steps. Step 1 aims to conduct a rapid scan of the entire design space with fewer designs, Step 2 focuses on increasing and confirming the approximation accuracy near the optimal design determined in Step 1. The proposed
sequential optimal design method can be applied to
design problems with regular and irregular boundaries
of design space by the benefit of extrapolation.

5) Rapid approximation of the entire design space
can be performed by presetting the threshold $d_{min}$ to be
larger, however, it can also feasibly be used to search
for relatively small clusters by adjusting the deviation
of random seeds and presetting the threshold $d_{min}$ to be
smaller in Step 1. However, the deviation of random
seeds is assigned larger values and the threshold ter-
mination is assigned smaller values in Step 1 to
increase the probability of solving the design within
the small cluster. In other words, a trade-off between
the accuracy over the entire design space and compu-
tation expense can be comprehensively achieved.

6) Application to the vehicle crashworthiness
design with multiple peaks reveals that the proposed
method is a feasible and practical approach.

Some points still need to be discussed, such as the
robustness of the optimal design and the confident
regions of estimated coefficients of holographic neural
network, which will be addressed in the future.

References

(1) Hagiwara, I., Review of Optimization Method and
Its Tendency, Forum on Spring Conf., Society of
Automotive Engineers of Japan, (in Japanese),
(1998 5).

(2) Proceedings of the 1st World Congress of Struc-
tural & Multidisciplinary Optimization, (1995),
Pergamon Press.

(3) Hagiwara, I., Innovative Technique and Future
Studies of Optimal Design from the Viewpoint
of Automotive Application Example, Proc. of the 7th
Design Engineering and Systems, (in Japanese),
No. 97-95 (1997), pp. 3-7, JSME.

(4) White, Jr, K.P., Simulation Optimization of the
Crashworthiness of a Passenger Vehicle in Fron-
tal Collision Using Response Surface Methodol-

(5) Giunta, A.A., Noise Aerodynamic Response and
Smooth Approximations in HSCT Design, AIAA-

(6) Roux, W.J., Stander, N. and Haftka, R.T.,
Response Surface Approximation for Structural
Optimization, AIAA-96-4042-CP, (1996), pp. 569-
578.

(7) Box, G.P. and Draper, N.R., Empirical Model-
& Sons, Inc.

(8) Carpenter, W.C. and Barthelemy, J.M., A Com-
parison of Approximations and Artificial Neural
Nets as Response Surface, AIAA-92-2247-CP,
(1992), pp. 2474-2482.

(9) Hagiwara, I., Shi, Q. and Kozukue, W., Develop-
ment of Non-linear Dynamic Optimization
Method Using the Holographic Neural Network,

(10) Shi, Q. and Hagiwara, I., Optimum Design Using
Holographic Neural Network and Its Extrapol-
ation, Proc. of the 7th Design Engineering and

(11) Hagiwara, I., Shi, Q. and Takashima, F., Develop-
ment of Crashworthiness Optimization Method
Using the Neural Network, (2nd Report: Ap-
Mech. Eng., (in Japanese), Vol. 64, No. 626, A

(12) Johnson, N.L. and Leone, F.C., Statistics and
Experimental Design in Engineering and the Phys-
ical Sciences, 2nd ed. Vols. 1 and 2, (1977), Wiley,
New York.

(13) LS-DYNA, (1997), Livermore Software Technol-
yogy Corporation

(14) Okamura, Y., Application Examples of Optimum
Design Tool PAM-OPT, Proc. of the 7th Design
pp. 451-454, JSME.

(15) Shi, Q., Hagiwara, I., Azetsu, S. and Ichikawa,
T., Optimization of Acoustics Problem Using Holo-
graphic Neural Network, Transaction of JSAE,