HDD SLIDER AIR BEARING DESIGN OPTIMIZATION USING A SURROGATE MODEL

Naozumi TSUDA
FUJITSU LABORATORIES LTD.
Atsugi-shi, Kanagawa, JAPAN

David B. Bogy
University of California at Berkeley
Berkeley, CA, USA

ABSTRACT
This report addresses a new optimization method in which the DIRECT algorithm is used in conjunction with a surrogate model. The DIRECT algorithm itself can find the global optimum with a high convergence rate. However the convergence rate can be much improved by coupling DIRECT with a surrogate model. The surrogate model known as the Kriging model is used in this research. It is determined by using sampling points generated by the DIRECT algorithm. This model expresses the shape of a hyper surface approximation of the cost function over the entire search space. Finding the optimum point on this hyper surface is very fast because it is not necessary to solve the time consuming air bearing equations. By using this optimum candidate as one of the DIRECT sampling points, we can eliminate many cost function evaluations. To illustrate the power of this approach we first present some simple optimization examples using known difficult functions. Then we determine the optimum design of a slider with 5nm flying height (FH) starting with a design that has a 7nm FH.

INTRODUCTION
In order to achieve a recording areal density of 1Tbit/in² in hard disk drives (HDD), the magnetic spacing between the read/write transducer and the magnetic medium must be reduced to approximately 5-6nm. As the FH of the air bearing slider is decreased, the ratio of the fly height modulation to the total fly height increases and degrades the read/write signal quality. Air bearing surface (ABS) designers have mainly focused on the static characteristics of the slider, such as FH, pitch and roll. However in order to design an ABS with small fly height modulation (FHM), the designer must consider the dynamic characteristics of the ABS. Dynamic simulation is much more expensive than static evaluation so the optimization of a slider design based on its dynamic characteristics becomes very expensive. Therefore we need to develop a searching algorithm that can find the global optimum design with a relatively small number of cost function evaluations.

PROPOSED SEARCHING METHOD
In optimization research in the field of DACE (design and analysis of computer experiments) [1] [2], the Kriging model is used as a surrogate model for the time consuming cost function. Usually a data sampling method is employed to construct the surrogate model. Some examples of data sampling methods for DACE are Monte Carlo sampling, Latin hypercube sampling [3], and orthogonal array sampling [4]. The data sampling method itself is unrelated to the function that is to be searched for the optimum. It just generates sampling points that are supposed to be suitable for constructing the surrogate model.

Data sampling methods for DACE tend to cover the entire search space whereas DOE methods tend to locate points at the boundary of the search space [4]. Since the DIRECT algorithm is a global optimization algorithm, it generates sampling points, as cost function evaluation points, that are spread over the entire search space. So we can use the DIRECT algorithm not only as the optimization algorithm but also as the data sampling algorithm.

There are two processes. One is the DIRECT process and the other is the Kriging process. DIRECT works as the optimization process using the ordinary DIRECT algorithm, except this process uses optimum candidates obtained from the Kriging process.

The Kriging process forms a Kriging model using sample points at which the cost function is evaluated in the DIRECT process. It is formed after each iteration of DIRECT. After forming the Kriging model, the optimum point on the Kriging model is searched for using the DIRECT algorithm version III (DIRECTIII) [5]. The MATLAB Kriging Toolbox [6] was used to form the Kriging model. We used 1st order polynomial and cubic type of correlation function for the Kriging model. After an optimum candidate is obtained in the Kriging process, this design is passed on the DIRECT process for further investigation. As this optimum candidate is obtained from the Kriging process, it does not have the cell information needed by the DIRECT algorithm for grouping and dividing cells. To
create these data, the cell that has the closest cost function value to the optimum candidate is searched for from the top of the line of the samples which is arranged in the order of its cell size and cost function value.

OPTIMIZATION RESULTS
To verify the proposed algorithm we used three simple functions as cost functions, which are defined as follows:
\[ f(x_i) = 100 \cdot (x_i - x_i^2)^2 + (1 - x_i)^2, \quad (x_i, x_j) \in [-2.048, 2.048] \]
\[ f(x_i) = \sum_{i=1}^{10} (x_i - 0.1 \cdot i)^2, \quad x_i \in [0,1], \quad i = 1, \ldots, 10. \]
\[ f(x_i) = \sum_{i=1}^{20} (x_i - 0.05 \cdot i)^2, \quad x_i \in [0,1], \quad i = 1, \ldots, 20. \]

We also tried to optimize slider ABS design. We chose the Information Storage Industry Consortium (INSIC) 7nm fly height Pico slider as the prototype slider (Pico refers to a certain set of overall dimensions of the slider). Its rail shape and constraint conditions are shown in Figure 1. The optimization problem defined here is to optimize the rail shape so that the flying height is 5nm and the roll profile is as flat as possible across the disk. The resulting number of design variables is eight. The disk rotation speed is 7600rpm. The air bearing simulations are performed at three radius positions designated OD, MD and ID. The air bearing simulator developed by Computer Mechanics Laboratory at UC Berkeley is used to solve static state of the air bearing. The largest weight of the cost function is set for flying height compared with for pitch and roll.

RESULTS AND DISCUSSION
Figures 2(a) ~ (c) show the optimization results. Figure 2(d) shows the optimization result of slider ABS design optimization. The solid marks represent the proposed method. The hollow marks represent the results using the DIRECTIII algorithm alone. We see the remarkable advantage in the 20-D case. All of these results show the DIRECT algorithm combined with the Kriging model converged faster than the DIRECTIII algorithms alone.

The Kriging model is a hyper-surface interpolated among the sample points. Therefore unexplored regions can be searched using the Kriging model without system analysis. Namely, a local-searching around a candidate region is done by using the Kriging model without fly height calculations. This is the main reason why the DIRECT algorithm combined with the Kriging model can get optimum designs faster than the DIRECTIII algorithms alone. However, the Kriging model does not have enough spatial resolution as the sample points generated by DIRECT algorithm. There is no way other than to wait until the DIRECT algorithm generates more samples to find a better design. We can see that the convergence rate suddenly improves after not improving after several iterations in Figure 2(a). It is conceivable that the DIRECT algorithm generated samples in the region where a better design might exist. By using these sample points the Kriging model could be formed with enough fidelity and a better design could be found without fly height calculations.

CONCLUSION
We combined the DIRECT algorithm and the Kriging model. This method enables us to find the global minimum faster than the DIRECTIII algorithm which is the fastest version among the DIRECT algorithms.