Preprint

This is the submitted version of a paper published in.

Citation for the original published paper (version of record):

Sun, X., Sellgren, U., Andersson, K. [Year unknown!]
Efficient and situated design of haptic devices.
[Journal name unknown!]

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-217326
Efficient and situated design of haptic devices

Xuan Sun∗
Department of Machine Design
KTH Royal Institute of Technology
Stockholm, 10044
Sweden
Email: xuans@kth.se

Ulf Sellgren
Associate Professor
Department of Machine Design
KTH Royal Institute of Technology
Stockholm, 10044
Sweden
Email: ulfse@kth.se

Kjell Andersson
Professor
Department of Machine Design
KTH Royal Institute of Technology
Stockholm, 10044
Sweden
Email: kan@kth.se

ABSTRACT

It is a challenging task to develop and optimize a high-performing haptic device for, e.g., medical training, because of the multi-objective performance requirements, and the complex relations between the design variables and the performance objectives. Consequently, the solution space is scrutinized with efficient optimization techniques, and the cause-effect relations are preferably represented in a way that enables further reasoning in a multi-discipline setting, e.g., as Pareto-front curves and surfaces. Furthermore, if we have found a Pareto-optimal solution, how can we benefit from the knowledge and optimization results obtained in that process if we are facing a trailing challenge (a new design situation) to re-design the initial solution to another application and/or other requirements, i.e. how can we address this new situation by efficiently re-using as much as possible of the initial solutions? In this paper, we first present multi-objective optimizations of a 6-degree-of-freedom TAU haptic device. To investigate how we efficiently can address new design situations, by making use of the initial solutions, we have defined, solved and analyzed six re-design cases. For the different cases, it is shown what can be reused and in what way that can be done. The six studied cases are then analyzed and a generic process for situated design optimization, based on available knowledge and Pareto-optimal solutions, is proposed.

1 Introduction

With the development of new technology, an increasing number of instruments are used in medical treatments. Haptic devices are one type of such tools, which are used, not only, in assisting surgery but also in training surgical skills in virtual reality. A haptic device is an actuated human-machine interface between human sense of touch and a real, tele-operated, or computer-generated virtual environment. Based on manipulation and interaction with objects in the predefined working environment, the haptic device can provide force and torque feedback to the operator [1]. Haptic devices to be used in medical applications, must be high-performing from several perspectives, which makes the design task quite complex, asking for efficient model-based and simulation-driven approaches. To enable simultaneous management of technical complexity

∗Corresponding author.
and optimized performance, the Multidisciplinary Design Optimization (MDO) [2] and the Metamodel-Based Design Optimization (MBDO) [3] methods have been integrated and implemented in a multi-tool framework [4][5]. The reasons for developing that framework is to make the design and optimization process of haptic devices more efficient and also flexible.

Most optimization examples published in the literature are focusing on well-defined design cases [4]. However, the design of haptic devices is fuzzy and complex [6] in the early phase, mainly due to the soft requirements related to touch and “feel”, a large solution space, significant nonlinearities, and the need for real-time estimation and control. This complexity can potentially be reduced by knowledge gained from simulations and/or physical tests. The knowledge gained in that process will, most likely, trigger new revisions of the principal design solution. These revisions may cause a significant amount of repetitive work and hence a much prolonged lead time. Consequently, in order to increase the efficiency of the design process, a significant challenge is to enable the knowledge gained in the process to be adapted and reused as much as possible, i.e. to be “situated” [7]. The notation of “situated” here indicates that the design is adaptable to new and/or changed situations which include a new/changed application and/or a modified/new set of performance requirements. A related challenge is to develop a methodology and a supporting framework to efficiently solve the revised or further developed design task by re-using the existing models, knowledge and results.

The long-term goal of the research presented in this paper is to further develop the previously proposed model-based framework and methodology into a situated and computationally efficient design framework for multi-objective optimization of haptic devices.

More specifically, we have addressed the following research questions:

- **R1:** What re-design scenarios may occur for haptic devices?
- **R2:** What knowledge gained in the optimization process of an initial design task potentially can be reused in future development scenarios?
- **R3:** How to efficiently solve the re-design scenarios?

The remainder of the paper is organized as follows: Section 2 describes the initial design case of a high-performing 6-degree-of-freedom (DOF) haptic device; Section 3 shows six hypothetical re-design scenarios and their solutions built on the initial design case. The studies of the situated design scenarios are concluded and analyzed in Sec. 4. In Sec. 5, the research in this paper and the limitation are discussed together with a generic re-design process graph. Concluding remarks and future work are provided in Sec. 6.

## 2 Initial design case

The presented study of the initial design and re-design situations is based on a 6-degree-of-freedom (DOF) TAU haptic device [8] to be used as a surgery training simulator. The TAU configuration (see Fig. 1) consists of a fixed I-column, a handle located on the Tool-centre-point (TCP) of a moving platform, and three kinematic chains which connect the platform to the fixed I-column. Chains 1 and 2 are symmetrical with a serial plus parallel linkage, while chain 3 is a pure serial linkage. Six motors are mounted on the I-column to enable the 6-DOF motion, and the rest of the joints are all passive.

### 2.1 System requirements

According to the general requirements for haptic devices, combined with the specific TAU simulator system requirements presented by Khan et al. [8], three of the listed preliminary requirements are used as requirements for the initial design task:
2.2 Performance criteria

Ahmad et al. [9] provided a list of the most commonly used performance criteria. Based on our original requirements, we judged the goodness of the haptic device from the two criteria: workspace volume and kinematic isotropy.

2.2.1 Workspace volume index (VI)

The workspace volume for the device is quantified by analyzing the dexterous workspace. The approach is to evaluate the dexterous workspace as proposed by Ahmad et al.[9]. In this approach, the prescribed rotations are applied at the TCP for each grid point in the workspace (see Fig. 2). Only the reachable points, which satisfies all constraints (i.e. singularity, limit of joints), are added to the dexterous workspace volume index (VI) given in Eq. (1). A binary flag $N$ (in Eq. (2)) is used to define if the grid point moving at the cross-section of the workspace is reachable or not.

$$\text{VI} = \int_{C_{\min}}^{C_{\max}} \int_{r_{\min}}^{r_{\max}} \int_{\theta_{\min}}^{\theta_{\max}} N(C_i, r, \theta) \cdot rdrd\theta$$

(1)

$N(C_i, r, \theta) = \begin{cases} 
1 & \text{if } P \in W(X, Y, Z, \alpha, \beta) \\
0 & \text{if } P \notin W(X, Y, Z, \alpha, \beta) 
\end{cases}$

(2)

$W$ represents grid points traversed by TCP in the $X$, $Y$ and $Z$ directions in the TCP local coordinate system $\{C\}$, with rotations along $X$-axis (pitch; $\alpha$), and $Y$-axis (yaw; $\beta$), and $Z$-axis (roll; $\gamma$). If the grid point $P$ is a reachable point in $W$, this point will be added to the workspace volume index. The range and step size of each parameter used to search the dexterous workspace is presented in Table 1, and the dimension of the maximum rectangular workspace volume ($W_x$, $W_y$, $W_{z_{\max}}$ and $W_{z_{\min}}$), as shown in Fig. 3, is searched within the dexterous workspace which should be not smaller than the minimum translational workspace. For simplification, only the grip point with the same value of $W_x$ and $W_y$ was picked for finding the maximum rectangle. Accordingly, the maximal rectangular workspace is $(2 \cdot W_y) \times (2 \cdot W_y) \times (|W_{z_{\min}}| + W_{z_{max}})$ [mm].

2.2.2 Global kinematic isotropy index (GII)

The second criterion, the kinematic isotropy, indicates how evenly the system moves in all six generalized directions in the dexterous workspace. The approach proposed by Gao and Gruver [10] is used for calculating the isotropy index $\text{II}$ in Eq. (3) where $\| J_m \|$ is the frame-invariant Euclidian norm of the Jacobian matrix [9].

$$\text{II} = \frac{1}{\| J_m \| \| J_m^{-1} \|}$$

(3)
Fig. 3. The dexterous and effective workspace footprint

Table 1. Parameters used in the workspace volume model

<table>
<thead>
<tr>
<th>Name</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>Step size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cx [mm]</td>
<td>-50</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>r [mm]</td>
<td>0</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>θ [°]</td>
<td>0</td>
<td>360</td>
<td>25</td>
</tr>
<tr>
<td>γ [°]</td>
<td>-10</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>α [°]</td>
<td>-10</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>β [°]</td>
<td>-10</td>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

The global kinematic isotropy index $GII$ in the dexterous workspace (in Eq. (4)) is defined as the average of the local kinematic isotropy $II_p$ in Eq. (5) within the dexterous workspace.

$$GII = \frac{\int_I II_p dv}{VI} \tag{4}$$

$$II_p = \frac{\sum_{e=0}^m \sum_{f=0}^n \sum_{g=0}^q II_{e,f,g} \alpha_\gamma \beta_\gamma}{m \times n \times q} \tag{5}$$

In Eq. (4), $dv$ is an infinitesimal volume element in the workspace, and $VI$ is the total volume of the workspace. In Eq. (5), $m$, $n$, and $q$ are the maximal rotations applied in roll ($\gamma$), pitch ($\alpha$) and yaw ($\beta$), respectively. The range of the global kinematic isotropy index $GII$ is from 0 to 1. When the value of $GII$ approaches to 1, the haptic device gives uniform motion in all workspace directions.

2.3 Problem formulation and optimization strategy

In the optimization problem, the target is to find an optimal mechanical configuration with maximum dexterous workspace volume index ($VI$) and global kinematic isotropy index ($GII$). Accordingly, five continuous structural design variables are varied during the optimization as shown in Fig. 4, where $\{G\}$ is the global coordinate system located on the fixed I-column and $\{C\}$ is the local coordinate system on the TCP. The five structural design variables are:

1. Shorter arm length ($L_1$) connected with three motors attached on the I-column,
2. Longer arm length ($L_2$) connecting the shorter arm with the moving platform (end-effector),
Fig. 4. Structural design variables for design of the TAU haptic device

3. Radius of the platform ($R_p$).

4. Angle of the attached joint ($\theta_p$) with chain 1 and chain 2 ($e_1$, $e_2$, $d_1$ and $d_2$ marked in Fig. 4) around the TCP local X-axis. Note that $L_3$ is the distance between each parallel chain, $L_3 = 2 \cdot R_p \cdot \sin(\theta_p)$.

5. The starting elevation angle ($\theta_6$) of chain 3 which defines the origin of the workspace.

MDO methods [2] are classified into single-level (SL) and multilevel (ML) methods. In a single-level method, only one optimizer is used to solve the design problem. The most common and basic single-level optimization method is the multidisciplinary feasible (MDF) method which enforces the multidisciplinary consistency for every set of design variables hence not suitable for strongly coupled problems [11]. And the other two SL methods, all-at-once (AAO) and individual-discipline-feasible (IDF), decouple the subspace analyzers to enable parallelism and check the multidisciplinary consistency in the system level hence more efficient and robust [12], especially for strongly coupled problems comparing to MDF, but more complex to use.

By utilizing ML methods, the complex design problem can be decomposed into a set of smaller sub-problems, and each sub-problem has an optimizer [13][14]. It was shown by Yi et al. [12] that the Bi-level Integrated System Synthesis (BLISS) method required a relatively small number of function calls to find a good optimum solution.

Based on the property of this initial design case, the MDF method was used to solve the stated optimization problem formulated in Eq. (6).

\[
\begin{align*}
\text{maximize} & \quad f_1(X) = VI(X) \\
\text{maximize} & \quad f_2(X, J) = GHI(X, J) \\
\text{over} & \quad X = [L_1, L_2, R_p, \theta_p, \theta_6] \\
\text{subject to} & \quad g_J(X) = \det(J) > 0 \\
& \quad g_V(x) = |W_Y(X), W_{z_{\text{min}}}(X), W_{z_{\text{max}}}(X)| \geq 25 \\
& \quad |W_X(X)| = |W_Y(X)| \\
& \quad X_{i_{\text{low}}} \leq X_i \leq X_{i_{\text{up}}}, i = 1, \ldots, 5
\end{align*}
\]

where $X$ is a vector of the five design variables, and $J$ is the Jacobian matrix. As constraints, the singularity condition ($g_J$) should not be equal or less than 0, and the four dimensions of the maximum rectangle ($W_x$, $W_Y$, $W_{z_{\text{min}}}$ and $W_{z_{\text{max}}}$) in the dexterous workspace should be not less than 25 in the TCP local coordinate system. The four dimensions indicate the distance between the rectangle corner and the initial TCP position in $X$-, $Y$- and $Z$-axis, as marked in Fig. 3. The initial design variable ranges are listed in Table 2.

Due to the complex character of the TAU mechanism, it was computationally exhaustive to evaluate the behavior of the full system model. To efficiently and effectively solve the design problem, the MBDO method was used in the design optimization process. A metamodel [15] is a surrogate model, i.e. a “model” of a model, for a computer-intensive analysis.

### Table 2. Initial design variable ranges

<table>
<thead>
<tr>
<th></th>
<th>$L_1$ [mm]</th>
<th>$L_2$ [mm]</th>
<th>$R_p$ [mm]</th>
<th>$\theta_p$ [$^\circ$]</th>
<th>$\theta_6$ [$^\circ$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>100</td>
<td>100</td>
<td>40</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>Lower</td>
<td>200</td>
<td>200</td>
<td>60</td>
<td>60</td>
<td>40</td>
</tr>
</tbody>
</table>
Table 3. Parameter values for optimization used in modeFRONTIER

<table>
<thead>
<tr>
<th>Scheduler</th>
<th>NSGA-II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of generation</td>
<td>100</td>
</tr>
<tr>
<td>Directional cross-over probability</td>
<td>0.5</td>
</tr>
<tr>
<td>Selection probability</td>
<td>0.05</td>
</tr>
<tr>
<td>Mutation probability</td>
<td>0.1</td>
</tr>
<tr>
<td>DNA string mutation ratio</td>
<td>0.05</td>
</tr>
<tr>
<td>DOE algorithm</td>
<td>ULHS</td>
</tr>
<tr>
<td>DOE number of designs</td>
<td>100</td>
</tr>
<tr>
<td>Total number of iterations</td>
<td>$100 \times 100 = 10000$</td>
</tr>
</tbody>
</table>

or simulation process. The main steps for performing MBDO are listed below. Several iterations might be needed within this process (referred as model-refinement) in order to get accurate metamodels and optimal solutions.

1. Perform Design of Experiment (DOE) [16] to study the effect of design variables and their interaction effects on the objectives, and potentially decrease the complexity of the design problem,
2. Construct metamodels of the system response factors using a set of data from sample experiments,
3. Validate the accuracy of the metamodels,
4. Solve the optimization problem using the constructed metamodels,
5. Validate the accuracy of the selected optimal solution on the full system model.

Based on the size of the design problem, we used a 2-level full factorial design [16] to perform the DOE study and hence to find the effects of the design variables on the two design objectives.

From previous studies of metamodeling techniques applied on TAU device [5], 300 sample points generated by Uniform Latin Hypercube sampling (ULHS) method [17], which generates randomly and uniformly distributed design points in each dimension, was suitable for constructing metamodels of the two objectives ($VI$ and $GII$). Furthermore, the most accurate metamodeling method of $VI$ and $GII$ was suggested to be Kriging (KR) [18] and Radial Basis Functions (RBF) [19], respectively. Hence, these findings were directly used in the MBDO process of the presented initial design case. The $k$-fold cross-validation method [20] was used to evaluate the error of the metamodel ($E_{kfold}$). The cross-validation process in the multi-tool environment followed the multi-tool framework proposed by Sun et al. [5].

For the constraint functions ($g_f$ and $g_v$), instead of constructing metamodels, the infeasible range of design variables could be filtered out during the DOE process. Therefore, the constraint functions could be ignored in the optimization process but checked when validating the optimal solution with the full model.

Due to the non-linear behavior of the TAU haptic device and the multi-objective character of the design problem, a general optimization process based on the Non-dominated Sorting Genetic Algorithm (NSGA-II) [21] was used to optimize the design solution. In our study, we used a software tool called modeFRONTIER® [22] to integrate different heterogeneous models and also automatically optimize the defined multi-objective and multidisciplinary design problem. The optimization scheduler in Table 3 was implemented in the modeFRONTIER framework. The best-fit metamodels of the two objectives constructed previously (KR for $VI$ and RBF for $GII$) were used in this optimization process.

As the final step, the optimal solution selected from the Pareto-front should be validated with the full model. Since the focus was not on the weighting of the objectives, 50 optimal designs were randomly picked and validated with the full model. The accuracy of the optimal solutions ($E$) shown in Eq. (7) was evaluated as the mean of relative absolute error (RAE in Eq. (8)), which should be no larger than 5%.

$$E = \frac{1}{n} \sum_{i=1}^{n} RAE_i \leq 0.05, n = 50$$ (7)

$$RAE_i = \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$ (8)

where,
Table 4. Main effect and significance of each factor on system response

<table>
<thead>
<tr>
<th></th>
<th>$L_1$</th>
<th>$L_2$</th>
<th>$R_p$</th>
<th>$\theta_p$</th>
<th>$\theta_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GII</td>
<td>-0.609</td>
<td>0.209</td>
<td>0.956</td>
<td>-0.406</td>
<td>-1.000</td>
</tr>
<tr>
<td>VI</td>
<td>0.843</td>
<td>-0.740</td>
<td>0.101</td>
<td>0.845</td>
<td>1.000</td>
</tr>
<tr>
<td>Wy</td>
<td>-0.091</td>
<td>1.000</td>
<td>-0.136</td>
<td>-0.636</td>
<td>-0.341</td>
</tr>
<tr>
<td>$W_{z_{\max}}$</td>
<td>0.585</td>
<td>-1.000</td>
<td>0.113</td>
<td>1.000</td>
<td>0.623</td>
</tr>
<tr>
<td>$W_{z_{\min}}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

([signi.]) $[0.048]$ $[0.292]$ $[0.005]$ $[0.138]$ $[0.003]$

$y_i$ = actual system response on the $i^{th}$ design point,

$\hat{y}_i$ = predicted system response on the $i^{th}$ design point.

2.4 Results

The effect and significance of each design variable on each objective and workspace constraint from the DOE are shown in Table 4. The effect value is between -1 to 1, and the significance varies from 0 to 0.5. If the effect value is close to unity, and the value of significance is close to 0, the design variable has a true and significant positive/negative effect on the response factor.

According to Table 4, except for design variable $R_p$, all design variables had contradictive effects on the two response factors, global isotropy index ($GII$) and volume index ($VI$). For example, $L_1$ had a negative effect on GII and a positive effect on $VI$. This indicated that the two response factors were in conflict with each other. Since the design variable $R_p$ had a significant positive effect on GII and less effect on $VI$, instead of being a design variable, $R_p$ was set to constant at its maximal value (60 mm). Furthermore, the data in the table also indicated that the design variables had no effect on the constraint $W_{z_{\min}}$, which remained a constant value of -50 [mm].

The range of design variables after DOE study (shown in Table 5) was used in the MBDO process. The solutions and its Pareto-front are illustrated in Fig. 5. The error of metamodels ($E^k$-fold) and the error of the Pareto-front ($E$) validated with the full model are listed in Table 6. According to the system response of the 50 validated optimal designs, the optimal values of each element in the workspace volume constraints ($g_V$) (illustrated in Fig. 3) are: $W_x=W_y=28$ or 35; $W_{z_{\min}}=-50$; $W_{z_{\max}}=25$, 30 or 35. One optimal solution and the corresponding design configuration from the Pareto-front are shown in Table 7 while its rectangular workspace was $70 \times 70 \times 75$ [mm].
Table 5. Range of design variables after DOE study

<table>
<thead>
<tr>
<th></th>
<th>L₁ [mm]</th>
<th>L₂ [mm]</th>
<th>Rₚ [mm]</th>
<th>θₚ [°]</th>
<th>θₜ [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>135</td>
<td>150</td>
<td>60</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>Lower</td>
<td>145</td>
<td>170</td>
<td>30</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

Table 6. Errors of metamodel and Pareto-front

<table>
<thead>
<tr>
<th>Validation process</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>k-fold cross-validation of metamodel VI</td>
<td>0.0239</td>
</tr>
<tr>
<td></td>
<td>GII 0.0102</td>
</tr>
<tr>
<td>Pareto-front validation with full model VI</td>
<td>0.0484</td>
</tr>
<tr>
<td></td>
<td>GII 0.0194</td>
</tr>
</tbody>
</table>

Table 7. One optimal design configuration for the initial design task

<table>
<thead>
<tr>
<th>Design variables</th>
<th>System response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Value</td>
</tr>
<tr>
<td>L₁</td>
<td>144.45 [mm]</td>
</tr>
<tr>
<td>L₂</td>
<td>150 [mm]</td>
</tr>
<tr>
<td>Rₚ</td>
<td>60 [mm]</td>
</tr>
<tr>
<td>θₚ</td>
<td>29.882 [°]</td>
</tr>
<tr>
<td>θₜ</td>
<td>27.105 [°]</td>
</tr>
</tbody>
</table>

3 Re-design cases

From the solution of the initial design case, we have a set of optimal designs which provide the values of global isotropy index (GII) varying from 0.4304 to 0.5228, and values of workspace volume index (VI) varying from 0.8939 to 0.3977 dm³. However, in the further product development process, or alternatively in a new development process, we might want to improve or change the device performance targets based on the gained knowledge of the system and new or modified system requirements. To address the new situation, the design optimization problem could be separated into six alternative re-design cases. To take the initial design case as the base design, the six cases, or new situations, can be characterized as:

1. Increased isotropy requirement: larger GII value,
2. Decreased isotropy requirement: smaller GII value,
3. Increased workspace volume: larger value of VI, W₁, W₂, |W₉min| and W₉max,
4. Decreased workspace volume: smaller value of VI, W₁, W₂, |W₉min| and W₉max,
5. Add new performance requirement(s): add performance criteria,
6. Remove obsolete requirement(s): remove performance criteria.

Re-design cases from 1 to 4 can be represented by a single design scenario, that is, the change is made of the value of existing objectives. The only difference is that, for cases 1 and 2, only the requirement of one objective is changed, but for 3 and 4, the limitation of the constraint functions are also changed. However, as stated in the initial design case, the infeasible range of design variables which unsatisfied the constraints was filtered out, and the constraint functions were ignored in the optimization process. Hence, the method to solve the re-design cases from 1 to 4 is the same. However, there are two possible situations that could occur in these four re-design cases, that is, the solutions satisfying the new requirements can be found in the initial (existing) optimal solutions, or if it is not present.

Another scenario is that we want to add a new requirement, which has not been considered for the existing initial design, or we want to design a haptic device to be used in another field, but with the same type of mechanism. This scenario can be classified into two re-design cases; either adds new requirements or removes less important requirements from the design case (re-design cases 5 and 6).

Solving the first four re-design cases represent two different situations. Examples of these re-design cases are elaborated on in the following subsections. In order to ignore repeated studies, the four cases were classified into two groups based on the different situations. Section 3.1 describes re-design cases 1 and 2, with the situation that the new requirements can be
satisfied with the existing solutions. Re-design cases 3 and 4, with another situation that no existing solution can satisfy the new requirements, are described in Sec. 3.2. Re-design cases 5 and 6 are described in Sec. 3.3 and Sec. 3.4, respectively.

3.1 Re-design cases #1 & #2: changed isotropy requirement

As an example of the revised system requirement, we want to have a device which satisfies the original requirements but also guarantees a relatively high isotropy performance. Since relatively high isotropy is an ambiguous requirement, we sharpen this requirement with an explicit limitation on the isotropy index, for example, the value of \( GII \) should be larger than 0.5.

To handle this new or sharpened requirement, the isotropy index was treated both as an objective and a constraint in the optimization process. In the Pareto-front for the initial design case (shown in Fig. 5), the range of the optimal isotropy index \( (GII) \) value was from 0.43 to 0.52 which contained solutions satisfying the new requirement. Hence instead of re-performing the optimization process from the beginning, solutions satisfying the requirement could be directly filtered from the Pareto-front and used as a solution for this re-design case. The filtered Pareto-front is shown in Fig. 6.

3.1.1 Verification study

Instead of using the direct-filtering method, the new design case was re-solved from scratch. The initial range of design variables was the same for the initial case (shown in Table 2). After the DOE study, the range of design variables was changed to as shown in Table 8. 300 sample points generated by ULHS method were used to construct metamodels of the performance criteria using the metamodeling methods KR and RBF for \( GII \) and \( VI \), respectively. The error of each metamodel was evaluated by the k-fold cross-validation method (shown in Table 9).

To illustrate the result of the verification study, all solutions and the Pareto-front are illustrated in Fig. 7(a). As a final step of this study, some optimal solutions from the Pareto-front were selected and validated on the full model (shown in Fig. 7(b)). The error of the selected solutions on both objectives (shown in Table 9) were less than 5% which satisfied the
accuracy requirement and hence could be used as optimal solutions of the re-design case.

To verify if the direct-filtering method can provide similar solutions to that from re-solved process, the Pareto-fronts filtered from the initial design and from the re-solved case were compared in Fig. 8. A comparison of the sensitivity of isotropy with respect to the variation of the design variables is illustrated in Fig. 9(a) - Fig. 9(d).

As shown in Fig. 8, the Pareto-front from the two different solving methods provides similar optimal values for the two objectives. By comparing the values of each design variable along the isotropy index (in Fig. 9), the difference between the optimal design configurations provided by the two different methods is small. This indicates that, instead of resolving the re-design case, the optimal solutions could be directly filtered from the existing Pareto-front, and hence, decrease the time of the re-design and improve efficiency.
3.2 Re-design case #3 & #4: changed workspace volume requirement

Consider the case that we want to complement our current haptic devices with another device with a larger workspace but also with as high isotropy as possible. In this product variant, we want a minimum workspace of 100x100x100 [mm] and highest possible isotropy. In this case, the requirements for existing performance criteria are changed and none of the existing solutions can be used to satisfy this new design case. Hence, the range of design variables should be changed and a rework should be done in the optimization process.

Since the performance criteria are kept the same as in the initial design case, the formulation of the optimization problem is the same as the initial design case in Eq. (6), only the constraint for the workspace volume \( g_V \) is changed. As the first step of the revision, the DOE results could help to find the new range of design variables based on the effect they had on the response factors as shown in Table 4. Thereafter the same MBDO steps as for the initial design case were performed.

The initial DOE results were part of the approach to efficiently adjust the design variables range in a re-design case presented by Sun et al. [23]. For more information about this approach and the underlying study, see Sun et al. [23].

3.3 Re-design case #5: add a new performance criterion

One of the most important system requirements is to provide realistic force feedback at the TCP. If the inertia at the TCP is too large, the structure of the device might give an unrealistic feedback to the user’s hand, with time-delays and vibrations. Therefore, besides the performance criteria listed in the initial design case, the inertia could be considered as a performance criterion, and hence to be defined as re-design case #5.

3.3.1 Kinematic inertia

For calculating the kinematic inertia, transformed to the TCP, instead of using the equation-based modeling method, the geometry-based model that is available in the MSC software Adams [24] was used. In the Adams model, the TCP was sequentially moved in the dexterous workspace found previously and the moment of inertia tensor in each point was recorded. Many points in the workspace could be used to evaluate the inertia performance. However, since the focus of this paper was to study how re-design cases could be made more situated and efficient, thirteen pre-defined points on the maximum rectangle were used for evaluating the inertia performance (marked in Fig. 10). Since our interest was in the moment of inertia about \( X, Y \) - and \( Z \) -axis, the other components of the moment of inertia tensor were ignored. The moment of inertia, as observed in the TCP, with respect to the global coordinate system \( \{G\} \) shown in Fig. 4 was stored as:

\[
I = [I_{xx} \quad I_{yy} \quad I_{zz}]
\]  

\( \quad \text{(9)} \)

Four indexes were used to evaluate the kinematic inertia, i.e., the sum of median values \( SI \) and variance of the moment of inertia \( \text{VMI} \) about the \( X, Y \) - and \( Z \) -axis shown in Eq. (10) and Eq. (11), where \( MI \) (in Eq. 12) represents the median value of the moment of inertia \( \text{II} \) at the TCP.

\[
SI = \sum MI_i, i = \{xx, yy, zz\}
\]

\( \quad \text{(10)} \)
the system-level and subsystem level optimization problem were formulated in Eq. (13) and Eq. (14), respectively. The isotropy performance. Based on these two targets, the moment of inertia at the TCP should be as low as possible. Hence, also decouples the coupling variables to enable parallel computing of the lower level problems. Furthermore, BLISS-2000 also utilizes the response surfaces to represent each discipline which enables minimal communication between disciplines and therefore potentially a higher efficiency.

3.3.2 Problem formulation and optimization strategy

In this re-design case, besides the two original objectives $GII$ and $VI$, four new criteria ($SI$, $VMI_{xx}$, $VMI_{yy}$ and $VMI_{zz}$), which evaluate the moment of inertia at the TCP, were added as performance objectives. The design variables remained the same as in the initial design case described in Sec. 2.3. Four constraints ($Wx$, $Wy$, $Wz_{max}$ and $Wz_{min}$) from the workspace volume model were also used as coupling variables [25] in the inertia model. The relations between variables, objectives and constraints are illustrated in Fig. 11.

Due to the computationally exhaustive and interactive system models, this re-design case was solved with one of the ML MDO method called BLISS-2000 [25]. This method decomposes the optimization task into upper and lower levels and further utilizes the response surfaces to represent each discipline which enables minimal communication between disciplines and therefore potentially a higher efficiency.

According to the new system requirement, the final design should first guarantee the workspace volume requirement and then subsequently the isotropy performance. Based on these two targets, the moment of inertia at the TCP should be as low as possible. Hence, the system-level and subsystem level optimization problem were formulated in Eq. (13) and Eq. (14), respectively.

The system-level optimization problem:

Given:

$Wx^a, Wy^a, Wz_{max}^a, M_{Ixx}^a, M_{Iyy}^a, M_{Izz}^a, VMI_{xx}^a, VMI_{yy}^a, VMI_{zz}^a$

Maximize:

$F_1(X) = GII^a(X)$
$F_2(X) = VI^a(X)$

Minimize:

$F_3(X, Y^*) = M_{Ixx}^a(X, Y^*) + M_{Iyy}^a(X, Y^*) + M_{Izz}^a(X, Y^*)$
$F_4(X, Y^*) = VMI_{xx}^a(X, Y^*)$
$F_5(X, Y^*) = VMI_{yy}^a(X, Y^*)$
$F_6(X, Y^*) = VMI_{zz}^a(X, Y^*)$

over $X = [L_1, L_2, R_p, \theta_p, \theta_\theta]$ $Y^* = [Wx^*, Wy^*, Wz_{max}^*, Wz_{min}^*]$

Subject to:

$Wx^*(X) - Wy^a(X) < 5$
$Wz_{max}^*(X) - Wz_{max}^a(X) < 5$
$Wz_{min}^*(X) - Wz_{min}^a(X) = 0$
$X_{low} \leq X_i \leq X_{up}, i = 1, \ldots, 5$

\[ VM_{Ii} = \frac{(l_{i1} - M_{Ii})^2 + \ldots + (l_{im} - M_{Ii})^2}{m}; \]
\[ i = \{xx, yy, zz\}; m = 13 \quad (11) \]

\[ M_{Ii} = \frac{\sum_{j=1}^{m} l_{i,j}}{m}; i = \{xx, yy, zz\}; m = 13 \quad (12) \]
The subsystem level optimization problem:

\[
\begin{align*}
\text{Maximize: } & f_1(X) = VI(X) \\
\text{Maximize: } & f_2(X, J) = GII(X, J) \\
\text{over } & X = [L_1, L_2, R_p, \theta_p, \theta_b] \\
\text{Subject to } & g_1(X) = \det(J) > 0 \\
& g_2(X) = |W_y(X), W_{z_{\text{max}}}(X), W_{z_{\text{min}}}(X)| \geq 25 \\
& |W_x(X)| = |W_z(X)| \\
& X_{\text{low}}^i \leq X_i \leq X_{\text{upp}}^i, i = 1, \ldots, 5 \\
\text{Find } & Y_o^\wedge = [VI_o^\wedge, GII_o^\wedge, W_y^\wedge, W_{z_{\text{max},o}}, W_{z_{\text{min},o}}] \\
\end{align*}
\]

where \( X \) = Vector of shared design variables affecting two or more modules, 
\( Y \) = Vector of system response (behavior variables), 
\((\cdot)\) = Metamodels of functions, 
\((\cdot)^a\) = Functions or variables input into a criterion from other modules, 
\((\cdot)^o\) = Functions or variables output from a module, 
\((\cdot)_{\wedge}\) = Optimal of functions or variables.

To solve this optimization problem efficiently, the existing knowledge from the initial design case was utilized. The optimization process is illustrated in Fig. 12. As subsystem level optimization problem, GII and VI should be optimized first which was already done in the initial design case. Hence, instead of performing re-work, the relevant data was imported from the database. Instead of using the full system model, metamodels of the functions and coupling variables were used for optimization. The range of design variables and the optimal values of the workspace volume \( (W_y, W_{z_{\text{max}}} \text{ and } W_{z_{\text{min}}}) \) from the initial design problem were used for constructing the metamodels of the inertia index and for generating the initial population in the system-level optimization process. The ULHS method was used here for generating the initial design points in each dimension.

In the optimization process, there were three coupling variables \( (W_y, W_{z_{\text{max}}} \text{ and } W_{z_{\text{min}}}) \), which were the output from the workspace volume model and the input into the inertia model. The consistency of these coupling variables was checked on the system-level. However, instead of a single optimal value, a set of integer values for \( W_y \) and \( W_{z_{\text{max}}} \) were found from subsystem level optimization. In order to check the consistency of these two on system-level, the metamodels of them were used to calculate an approximate value as a response to the shared design variables. Since the metamodel can only provide continuous function response, it was difficult to get these two metamodels accurate and to fit an integer value. But they could provide a system response within a range near the actual system response. Based on the optimal value of \( W_y \) and \( W_{z_{\text{max}}} \) found in the initial design case (Sec. 2.4), we defined that if the difference between the values of the input \( (W_y^\ast \text{ and } W_{z_{\text{max}}}^\ast) \) and the predicted output \( (W_y^\wedge \text{ and } W_{z_{\text{max}}}^\wedge) \) was smaller than 5, they were considered as consistent. For \( W_{z_{\text{min}}} \), since single optimal value was found previously, \( W_{z_{\text{min}}}^\wedge \) should be always equal to \( W_{z_{\text{min},o}} \), which is -50 [mm].

The same optimization scheduler shown in Table 3 was used to solve both the system and subsystem level optimization problems. As a final step, the selected optimal solution should be validated with the full model, e.g., with an error less than 5%. If the accuracy was not sufficient, a revision of some steps might be needed.

3.3.3 Results

The metamodels of the inertia criteria as well as the workspace volume constraints were constructed in parallel within the range of design variables and the optimal values of the coupling variables. As result, eight metamodels were constructed and used in a system-level optimization process. The corresponding information on each metamodel is shown in Table 10.

As results of the entire optimization process, sets of optimal data were chosen as the Pareto-front, from which the final optimal design could be selected. Since many objectives should be considered when selecting the final optimal design, the optimal values of the three objectives \( (GII, VI \text{ and } SI) \) were first cross-compared, as shown in Fig. 13. It was obvious that objectives GII and VI had opposite effects on each other. However, it was difficult to observe a clear effect relation between GII and SI, as well as between VI and SI. A bubble chart, which more clearly illustrates the correlation between the three objectives, is plotted in Fig. 14. Over the entire Pareto-front, we could found more than one design point with similar values of GII and VI but with different values of SI. In this case, the most optimal design point within the most relevant range of GII and VI was always the point with the lowest value of SI, as illustrated by the black curve in Fig. 14. Hence, based on the optimization results for these three objectives, we can select an optimal solution from this black curve based on the weighting of GII and VI.

However, before selecting the optimal design, the variance of each moment of inertia tensor \( (I_{xx}, I_{yy} \text{ and } I_{zz}) \) should preferably also be considered. A smaller variance indicates a more robust design, i.e., a design that is insensitive to variations. Fig. 15 illustrates the variance \( (VMI) \) with respect to the mean of moment of inertia \( (MI) \) about the X-, Y- and Z-axis,
respectively. As can be observed, the variance increased with the $MI$ value. However, the difference between the maximum and minimum values of the variance was rather small, e.g., the range of $VMI_{xx}$ was from $4.75$ to $5.3\ [kg^2 \cdot dm^4]$. This indicates that the design points in the Pareto-front gave a relative robust moment of inertia. Therefore, these variance indexes had less importance to the problem and could consequently be ignored.

At last, after a validation step, one optimal design marked by a red circle in Fig. 14 was chosen as the final optimal design. The design configuration of this optimal solution and its actual values for the design variables and the objectives are shown in Table 11. Its maximal rectangular workspace was $70 \times 70 \times 80\ [mm]$.

### 3.4 Re-design case #6: remove a performance criterion

Another scenario could also be to design a new series of haptic devices with the same mechanism as an existing product, but for a different application. For example, instead of a simulator for medical operation, a series of TAU-based haptic gaming games may be required. Under this condition, a low-cost haptic device with a large workspace may be the design targets. Hence some less important performance criteria, such as the inertia and isotropy indexes, can potentially be removed from the design requirements.

For such a re-design case, the correlation between the existing criteria defines the most suitable solving method, from the effectiveness point of view. If the correlation between the removed criteria and the remaining criterion/criteria is low, such as the correlation between the inertia index ($SI$) and the workspace volume index ($VI$), the data related to $SI$ can be directly filtered out, and the Pareto-front from the remaining solutions then becomes the new solutions to the re-design case.
Table 10. Construction information of the 8 required metamodels

<table>
<thead>
<tr>
<th>Function</th>
<th>Sample Size</th>
<th>Method</th>
<th>$E^{k-fold}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIxx</td>
<td></td>
<td></td>
<td>1.08e-5</td>
</tr>
<tr>
<td>MIyy</td>
<td></td>
<td></td>
<td>1.12e-5</td>
</tr>
<tr>
<td>MIzz</td>
<td>84</td>
<td>RBF</td>
<td>9.16e-6</td>
</tr>
<tr>
<td>VMIxx</td>
<td></td>
<td></td>
<td>7.76e-5</td>
</tr>
<tr>
<td>VMIyy</td>
<td></td>
<td></td>
<td>8.86e-5</td>
</tr>
<tr>
<td>VMIzz</td>
<td></td>
<td></td>
<td>1.05e-4</td>
</tr>
<tr>
<td>Wy</td>
<td>300</td>
<td>RBF</td>
<td>0.058</td>
</tr>
<tr>
<td>$W_{z_{max}}$</td>
<td>300</td>
<td>RBF</td>
<td>0.241</td>
</tr>
</tbody>
</table>

Fig. 13. Pareto-front cross-comparison: (a) GII vs. VI, (b) GII vs. SI, and (c) VI vs. SI

Fig. 14. Bubble chart of GII, VI and SI in the Pareto-front

However, if the correlation is high, such as a high correlation between $GII$ and $VI$ due to an opposite effect to each other, there might be higher $VI$ values with extreme low $GII$ values, which are not shown in the initial optimum solution space. A solution to this situation is to expand the design variable ranges, based on the effects of design variables on system responses (such as Table 4), and repeating the MBDO process from the DOE activity step.

4 Analysis of the situated scenarios

According to the studies of the six listed re-design cases, which represent new or modified applications and thus new requirements, i.e. new design situations, they can be categorized into three groups of design optimization cases:

1. Change the range of constraints or add new constraints,
2. Add new performance criteria,
3. Remove less important performance criteria.
To perform a re-design case, some rework can hardly be avoided. However, utilizing existing optimization solutions of the initial design case and knowledge of the system can potentially significantly reduce the lead time for the new design case. Besides the group of the changed optimization cases, the range of feasible solutions also varies with different re-design situations.

The re-design situations and their corresponding solutions are:

For group #1:
- If the new design case can be satisfied by the existing optimal solutions, new solutions can be obtained with filtering of the existing solution space for the initial design case;
- If no solution for the initial design case can satisfy the new re-design case, the current range of the design variables cannot provide feasible solutions. Hence the ranges of the design variables need to be expanded, and the design optimization process must be repeated. But, the DOE study and the effect table from the initial design case (such as Table 4) can be used to guide the process to find the new ranges of the design variables.

For group #2:
- If only local and coupling variables are added and the design variables are unchanged, the BLISS-2000 method can be used to solve this new optimization problem; The initial design optimization problem can then be treated as optimization on a subsystem level;
- If a re-design optimization problem requires new (shared) design variables added to both the initial and the new criteria, re-work of the entire design and optimization process is needed.

For group #3:
- If a removed performance criterion is shown to highly interact with other criteria, the range of the design variables might have to be expanded /changed, and the optimization process should, consequently, be re-analyzed, in order to assure a high performance on the remaining criteria. The results of DOE study from the initial design case can be utilized;
5 Discussion

An initial case with the task to design a high performing 6-DOF haptic device with two objectives was first studied. Six possible trailing re-design cases were then identified, described and situately elaborated on from an optimization efficiency point of view. During the studies of the re-design examples, some existing knowledge and information from the initial design case were found to be re-usable. Reusing information and results from a previous design project can potentially significantly reduce the lead-time of the new product.

The six re-design cases and the related situations were categorized, and their corresponding solving methods were merged into three distinct optimization groups that can be solved in three corresponding process blocks. The three design optimization blocks and their corresponding situated design scenarios are represented in Fig. 16 as a situated and generic process graph. In this figure, the gray blocks indicate the preferred process steps to find the set of solutions for different types of re-design cases, and the number marked below the action (①②③④) shows the existing knowledge which can be utilized in the situated re-design process. The re-design cases elaborated on in this paper are marked with the green background color.

All studied cases were multi-objective optimization (MOO) problems. However, the MDO methods, such as MDF and BLISS-2000, were used to solve the MOO problems here. Instead of a single optimal solution, a Pareto-front curve or surface was found from the optimization process using MDO methods, and a weighting of objectives could be performed afterward. This means that design decisions early in the process when there is large uncertainties and fuzziness can be postponed, which adds flexibility to the process. Or in other words, it makes the process more adaptable to new knowledge and situations, and thus more situated.

In re-design case #5, the inertia performance of the device was added as a new design optimization criterion. To evaluate the moment of inertia observed at the TCP, 13 pre-defined locations within the maximum rectangle in the dexterous workspace were used. In order to get a more refined estimate of the inertia performance of the device, more points within the dexterous workspace could be defined and evaluated.

Some of the situated scenarios have not been verified in the presented research. Furthermore, it is also possible to find more complex re-design cases than the ones studied, e.g. a combination of two re-design cases. The solutions to this kind of cases may depend on many other situations, which cannot be solved in a single action. Hence, further analysis is needed to structure and solve more complex cases than those studied here.

6 Conclusions and future work

The presented research identified and elaborated on six different 6-DOF haptic device re-design optimization cases that may occur after an initial design optimization task had been solved. Three groups of optimization cases were categorized, i.e. adding criteria, removing criteria, and modifying existing criteria. A process graph was proposed for conducting trailing re-design optimization tasks. The knowledge and information created and stored when optimizing an initial design case, that can enable re-design efficiency and situatedness are:
- The results from the DOE study and the table of design variable effects on system response,
- The ranges of design variables,
- The Pareto-front,
- The Metamodels of the existing criteria.

Three methods can be used to efficiently solve the re-design cases represented in the nested process graph:

- If the Pareto-optimal solutions for the initial design case satisfy a new re-design case, the new solutions can be directly filtered from the existing Pareto-front,
- If new criteria have to be added to the initial design optimization case and no changes of the shared design variables are made, the initial design optimization problem can be used as a subsystem level problem, and the entire problem can be solved with the BLISS-2000 method,
- For other cases, that require the design variable ranges to be expanded compared to the initial design case, the MBDO process must be re-executed. The new design variable ranges can be changed based on the results of the initial DOE study.

To further generalize the conclusions of the presented research, the following research tasks are planned for:

- Verify all re-design scenarios shown in the proposed situated framework,
- Identify and analyze how to efficiently address other potential re-design scenarios, which represent new, modified or aggregated situations.

Acknowledgment

This research was supported by the China Scholarship Council. We also would like to thanks the support from Enginsoft and MSCsoftware.

References