Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Knowledge Extraction from Aerodynamic Design Data and its Application to 3D Turbine Blade Geometries

Lars Gräning

2009

Preprint:

This is an accepted article published in International Workshop on Machine Learning for Aerospace. The final authenticated version is available online at: https://doi.org/[DOI not available]

Knowledge Extraction from Aerodynamic Design Data and its Application to 3D Turbine Blade Geometries

Lars Graening
Honda Research Institute Europe GmbH
Carl-Legien-Strasse 30
D-63073 Offenbach/Main, Germany
lars.graening@honda-ri.de

Abstract

Applying numerical optimization methods in the field of aerodynamic design optimization normally leads to a huge amount of heterogeneous design data. While usually often only the promising results are investigated and incorporated to drive further optimizations, general methods for investigating the entire design data set are rare. We propose methods that allow the extraction of comprehensible knowledge from design data represented by discrete unstructured surface meshes. The knowledge is prepared in a way so that it is usable for guiding further computational as well as manual design and optimization processes. A displacement measure is suggested in order to investigate local differences between designs. This measure provides information on the amount and direction of surface modifications. Using the displacement data in conjunction with statistical methods and data mining techniques provides meaningful knowledge from the data set at hand. The theoretical concepts have been applied to a data set of 3D turbine stator blade geometries.

1 Introduction

In aerodynamic design optimization the main goal is to find three-dimensional shapes, that are optimal for specific performance measurements, like aerodynamic drag or lift, under specific constraints, e.g. manufacturing limitations. In general, during the optimization process a large number of designs is generated and evaluated based on different geometric representations and parametrization. The results are heterogeneous design data sets from which only a very small number of designs are usually processed further, e.g. in rapid prototyping devices for experiments or for analyzing its detailed design characteristic. However, a lot of information about the process and the problem at hand is hidden in all of the data and can be condensed into comprehensive rules. We aim at exploiting this comprehensible knowledge which is contained in large design data sets. In order to be able to investigate the entire data set an universal representation of the designs is required. We suggest the use of unstructured surface meshes to represent the surface of the designs. In a pre-processing step the absolute information of the design surface is transformed into displacement data which together with statistical and data mining methods is used in order to extract meaningful knowledge from the designs and its performance numbers. The knowledge can be prepared in such a way that it is on the one hand usable by the engineer and on the other hand by a follow-up computational design and optimization process.

2 Universal Design Representation and Displacement Measurement

Different geometric representations make it difficult or even impossible to analyze the whole data set. Therefore, it is necessary to find an adequate representation that captures all shape variations and that can be applied to various data mining techniques. We suggest the use of unstructured surface meshes as a universal representation of the design surface. The unstructured surface mesh $\mathcal{M}$ is a partially linear approximation of the contour of the design. Each mesh consists of a list of vertices $\mathcal{V} = (\vec{v}_1, ..., \vec{v}_n)$, a list of polygons $\mathcal{K} = \{\{i_1, i_2, i_3, ..., i_\mu\}_k\}, k \in [1...m]$ and a list of normal vectors $\mathcal{N} = (\vec{n}_1, ..., \vec{n}_n)$. A vertex can be seen as a sample point of the contour of the design. The polygonal faces define the neighborhood relation between the vertices with i_l being the index of a vertex l. Each normal vector $\vec{n}_i$ has a defined direction perpendicular to the surface mesh and provides local curvature information at the position of vertex $\vec{v}_i$. As long as all necessary geometric representations can be transferred into unstructured surface meshes, this representation allows the analysis of local shape modifications and their influence on the performance value(s) independent of the parametric representation that has been used during the design and optimization process. For instance the majority of the CAD software allows the export of solids into stereo-lithography (STL) files which describe the designs as triangular unstructured surface meshes.

Given the surface mesh representation, the naive approach would be to analyze the absolute coordinates of the vertices in order to extract information on the design modifications. Instead of analyzing the absolute coordinates in this work it is suggested to calculate the displacement measures that quantify the local deformation between two design. The displacement is calculated between corresponding vertices of two surface meshes. This requires that a good estimation of the corresponding vertex does exist. More formally the displacement measure is defined as follows:

$$\delta_{i,j}^{r,m} = \delta(\vec{v}_i^r, \vec{v}_j^m) = (\vec{v}_j^m - \vec{v}_i^r) \circ \vec{n}_i^r, \delta \in (-\infty, +\infty), \tag{1}$$

and is the projection of the difference vector $\vec{s}_{ij} = (\vec{v}_i^r - \vec{v}_j^m)$ onto the normal vector $\vec{n}_i^r$ of vertex $\vec{v}_i^r$ of the reference design $\mathcal{M}_r$. The absolute value of the displacement measure provides information on the amount of vertex modification while the sign of the displacement measure in conjunction with the normal vector of the vertex provides information on the direction of the vertex modification. The normal vector $\vec{n}_i^r$ points towards the normal or positive direction of vertex modification.

Calculating the displacement quantities instead of using the absolute coordinates of the vertices provides certain advantages. One advantage is that the calculation of the displacement values leads to a reduction in the number of parameters that have to be handled by the knowledge extraction algorithm. Furthermore, because only one parameter value is assigned to each vertex the visualization and interpretation of the results becomes straightforward. More detailed information on the design representation, the displacement measurement and its properties can be found in [1,2].

3 Knowledge Extraction from Design Data

The calculated displacement values and the performance differences are the basis for the extraction of knowledge from the design data. This work provides means for extracting information on the modifications in the design space, its relation to the performance number(s) and the modeling of interrelated design modifications and its joint influence on the performance. To show the feasibility and the practical relevance, a few results from applying these techniques to a data set of a 3D stator blade from a Honda gas turbine [3] are presented.

3.1 Investigation of Design Deformations and Sensitivity

Analyzing local modifications in form of vertex displacements helps to gain some insight into the applied design modifications. Information on the differences between two designs is directly provided by the displacement measure. Two measures are suggested: First, the relative mean vertex displacement that provides information on how vertices have been modified by means of analyzing the displacement for all designs related to one reference design (Eq. 2) and second, the overall displacement variance that identifies the vertices that have been modified most frequently (Eq. 3).

$$\overline{\delta}_i^r = \frac{1}{N-1} \sum_{m=1, m\neq r}^{N} \delta_{i,j}^{r,m} \tag{2}$$

$$\sigma_{\delta_i} \approx \sqrt{\frac{2}{N(N-1)} \cdot \sum_{r=1}^{N} \sum_{m=r+1}^{N} (\delta_{i,j}^{r,m})^2} \tag{3}$$

These quantities provide knowledge about the design space and the explored design modifications. Potentially, this information can lead the aerodynamic engineer to new search directions and design concepts.

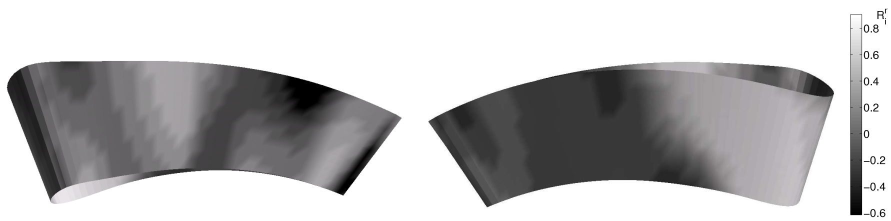

Figure 1: Illustration of the results of the sensitivity analysis mapped as gray values onto the blade surface. Left: shows the pressure side; Right: shows the suction side of the blade.

While Eq. 2 and Eq. 3 provide information on the design space, sensitivity analysis is applied to relate the displacement values to variations of the corresponding performance values. In order to identify vertices that are sensitive to performance changes based on the whole data set, the Pearson correlation coefficient is calculated based on all pairwise design comparisons:

$$R_i^r = \frac{\sum_{m=1, m\neq r}^{N} (\delta_{i,j}^{r,m} - \overline{\delta}_i^r)(\phi^{r,m} - \overline{\phi}^r)}{(N-1)\sigma_{\delta_i^r} \sigma_{\phi^r}} \tag{4}$$

where $\phi^{r,m} = f^m - f^r$ is the performance difference between two designs r and m, $\overline{\phi}^r$ is the mean value of the performance differences with respect to the reference design r. σ is the standard deviation of the displacements δ_i and ϕ respectively. Fig. 1 shows the result of the sensitivity analysis applied to a set of 200 blade designs, where the aerodynamic pressure loss defines the performance of each design. Brighter regions code for a positive correlation coefficient and darker regions for a negative one. Design regions with a positive (negative) correlation coefficient indicate that a deformation of this region to the outside of the blade will increase (decrease) the pressure loss (the normal vectors of the vertices point towards the outside of the blade).

3.2 Modeling and Analyzing Interrelated Deformations

For the calculation of the sensitivity described above, the displacement of each vertex is considered independent of the others. Especially in aerodynamics, the interrelation between distant vertices or design regions and their joint influence on the performance play an important role. In this section, special characteristics for the extraction of knowledge in form of associative rules based on data from unstructured surface meshes are discussed. The rules describe the relation between the displacement of distant vertices and their joint influence on the performance criteria.

In general, the number of input parameters must be kept small for most modeling techniques in order to produce a small set of interpretable and manageable association rules. Concerning displacement data, the number of inputs equals the number of vertices n, which is large in practice. Therefore, a reduction of the number of input parameters is strongly required. In order to reduce the number of parameters it is reasonable to combine neighboring vertices to form n_c design regions where $n_c \ll n$. During a design or optimization process, it is unlikely that only single vertices are modified. Rather entire design regions of a certain extent are considered for shaping new designs.

Our suggested procedure for dimension reduction by means of identifying local design regions is as follows. In a first step, the vertex sensitivity is calculated (Eq. 4). Vertices with a small sensitivity that do not seem to contribute on the performance are filtered out from the entire set of vertices by applying a threshold τ to the sensitivity value. Each of the remaining vertices is assigned to one of two sets $\mathcal{R}_+$ and $\mathcal{R}_-$. $\mathcal{R}_+$ contains vertices where $R_i > +\tau$ while $\mathcal{R}_-$ contains vertices with a correlation value $R_i < -\tau$. Finally, in order to form the desired design regions, a KMeans clustering algorithm is applied to each of the two resulting sets of vertices. The clustering is splitting the sets into clusters based on a predefined distance measure. Distant vertices are assigned to separate clusters and neighboring vertices to one and the same cluster. The gap statistic is used to overcome the problem of selecting an appropriate number of clusters in advance. Once the design regions have been identified the vertices closest to the cluster centers are considered for modeling and for the extraction of design rules. Fig. 2 shows the emerged cluster centers after applying the above algorithm to the blade data set.

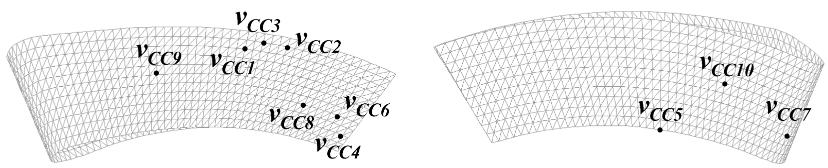

Figure 2: Illustration of the results from dimensionality reduction. The vertices closest to the cluster centers, resulting from the clustering of the filtered vertices are shown on the pressure side (left) and the suction side (right).

Rule induction is one of the fundamental and most often applied tools in the field of data mining and machine learning. Rules are easy to interpret by the engineer and hence raise his/her understanding of the system at hand. In aerodynamics the influence of one region of the design on the performance often strongly depends on the shape of the remaining design regions. Our driving force is to extract knowledge describing the complex relation between design regions and their performance number(s). It is important that the aerodynamic engineer is able to use the rules for the further development of new designs. In the present framework classification tree techniques are applied for rule extraction. This technique easily allows to control the number and complexity of the design rules. Design rules that are generated based on displacement data describe the interrelation of vertices and their influence on the performance relative to a predefined reference design. An example of a relative design rule from the blade data set is as follows:

$$IF \quad \delta_{CC8}^{r,m} > 0 \quad AND \quad \delta_{CC7}^{r,m} < 0 \quad THEN \quad \phi^{r,m} > 0$$

From this rule it is expected that a modification of the vertex $\vec{v}_{CC8}$ towards its normal direction in conjunction with a modification of vertex $\vec{v}_{CC7}$ against the direction of the normal vector will result in an increase of the pressure loss. It can be seen that the identified rules can easily be interpreted by an aerodynamic engineer.

In [2] the reliability of the extracted knowledge has been proven and this work provides additional information on the presented algorithms and results.

References

[1] L.Graening, M.Olhofer, B.Sendhoff, Knowledge Extraction from Unstructured Surface Meshes, In *Proceedings of 8th International Conference on Intelligent Data Engineering and Automated Learning*, IDEAL 2007, pp. 497-506, 2007

[2] L. Graening, S.Menzel, M.Hasenjaeger, T.Bihrer, M.Olhofer, B.Sendhoff, Knowledge Extraction from Aerodynamic Design Data and its Application to 3D Turbine Blade Geometries, *Journal of Mathematical Modeling and Algorithms*, 7:329-350, 2008

[3] M.Hasenjaeger, B.Sendhoff, T.Sonoda, T.Arima, Three Dimensional Aerodynamic Design Optimisation Using Single and Multi-Objective Approaches, Evolutionary and Deterministic Methods for Design, *Optimization and Control with Applications to Industrial and Societal Problems*, EUROGEN, Munich, 2005