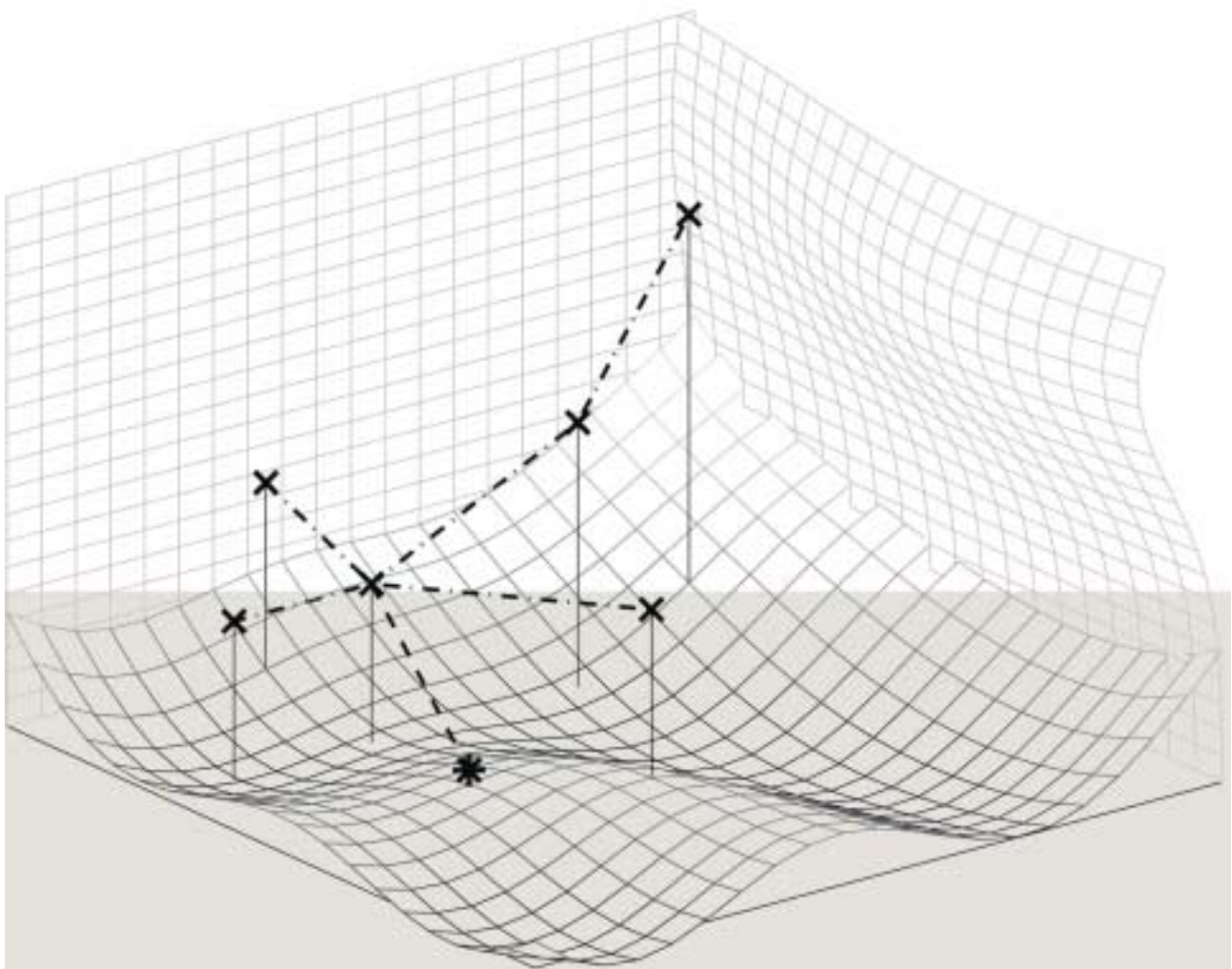


Modellbasierte Online-Optimierung moderner Verbrennungsmotoren

Teil 1: Aktives Lernen

Dieser zweiteilige Beitrag stellt den modellbasierten Optimierungsalgorithmus „mbminimize“ vor, der in Kooperation von der Universität Tübingen und der BMW Group für die Online-Optimierung von Verbrennungsmotoren entwickelt wurde. Der vorliegende erste Teil konzentriert sich auf das grundlegende algorithmische Design, auf Modellierung, Versuchsplanung und aktives Lernen. Der zweite Teil diskutiert Strategien zur Behandlung von Limits wie Motorklopfen.



1 Einleitung

Hersteller von modernen Verbrennungsmotoren sehen sich mit wachsenden Herausforderungen konfrontiert: Die gesetzlichen Umweltauflagen werden strenger, Kraftstoffpreise steigen, und die Kunden fordern gute Leistung und hohen Fahrkomfort bei geringem Verbrauch. Als Konsequenz wächst die Komplexität der Motoren stark, die Zahl der einstellbaren Parameter wird größer. Die Applikation des Motors wird so zu einem hochdimensionalen Optimierungsproblem, das den Einsatz rechnergestützter Verfahren erfordert.

Die bereits weitgehend etablierte Form der rechnergestützten Motoroptimierung ist die Offline-Optimierung. Derartige Verfahren werden etwa in [1], [2] oder [3] beschrieben und arbeiten prinzipiell wie folgt: Anhand eines Versuchsplans werden hinreichend viele Messungen am Motor vorgenommen. Aus diesen und ggf. geeignetem Vorwissen über den Motor wird am Rechner ein Modell erstellt, das das Motorverhalten simuliert. Nun kann das Modell am Rechner optimiert werden. Die gewonnenen Optima werden am Prüfstand verifiziert und anschließend zur Applikation verwendet.

Bei der Online-Optimierung interagiert das Optimierungssystem direkt mit dem Prüfstand, sodass die Applikation im Idealfall vollautomatisch abläuft. Diese Vorgehensweise bietet außerdem den Vorteil, dass die durch Messung gewonnenen Informationen sofort in den weiteren Verlauf der Optimierung einfließen können. Hierauf wird in den Abschnitten 3 und 4

ausführlich eingegangen. Vorteile der Offline-Optimierung sind dagegen, dass die Rechenzeit von der Prüfstandszeit getrennt werden kann, und dass der Prüfstand parallel zur Berechnung weiterhin manuell gefahren werden kann. Die automatische Prüfstandssteuerung mit Limitbehandlung stellt die Algorithmen vor neue Herausforderungen, dies ist Gegenstand von Teil 2 dieses Beitrags.

Auch Online-Optimierungssysteme sind bereits verfügbar, beispielsweise Cameo [4] und Vega [5]. Im vorliegenden Bericht wird ein weiterer Algorithmus für die Online-Optimierung, *mbminimize*, vorgestellt. Er unterscheidet sich von den bekannten in mehrfacher Hinsicht:

- **Modellierung:** Es wird ein globales Modell anstelle von einzelnen Modellen für jeden Betriebspunkt verwendet. Dadurch kann durch Messung gewonnene Information für mehrere Betriebspunkte genutzt werden, und der Messaufwand wird reduziert.

- **Modelltypen:** Es werden heterogene Komitees aus neuronalen Netzen und kubischen Regressionsmodellen eingesetzt.

- **Versuchsplanung:** Das Modell selbst bestimmt den Versuchsplan (aktives Lernen), die Zielgröße fließt zunehmend in den Versuchsplan ein.

- **Limitbehandlung:** Hier werden neue Strategien entwickelt, die einerseits gewährleisten, dass verbotene Bereiche in geeigneter Form modelliert werden können und das andererseits der Motor in einen sicheren Bereich gefahren wird, wenn ein Limit auftritt (siehe Teil 2 dieses Beitrags).

Die Autoren



Dr. Jan Poland war wissenschaftlicher Angestellter am Lehrstuhl für Rechnerarchitektur der Universität Tübingen.



Dipl.-Phys. Kosmas Knödler war wissenschaftlicher Angestellter am Lehrstuhl für Rechnerarchitektur der Universität Tübingen.



Dipl.-Ing. Thomas Fleischhauer arbeitet als Ingenieur und Algorithmenentwickler bei der BMW Group.



Dr.-Ing. Alexander Mitterer arbeitet als Ingenieur und Algorithmenentwickler bei der BMW Group.



Dipl.-Ing. Stephan Ullmann ist Doktorand bei der BMW Group im Bereich Motorentwicklung.



Prof. Dr. Andreas Zell leitet den Lehrstuhl für Rechnerarchitektur der Universität Tübingen.

... weniger Experimente
... mehr Effizienz!

Produktentwicklung mit
Fluent CFD-Software

Wer an der Spitze bleiben will, muss Entwicklungszyklen straffen und **teure Experimente reduzieren**. Aufgaben, die sich mit der Strömungssimulation (CFD) ideal lösen lassen. FLUENT, weltweit führend im CFD Markt, bietet die Software dazu.

Einfach zu bedienen, aber mit ungeahnten Möglichkeiten in Modellierungsumfang, Realitätstreue und Flexibilität. Plus einen kompletten Service rund um die CFD-Technologie. **So hilft Ihnen FLUENT, Zeit und Kosten zu sparen.**

 **FLUENT**

www.fluent.de

Birkenweg 14a · D-64295 Darmstadt · Tel. +49 (0)61 51 36 44-0 · Fax +49 (0)61 51 36 44-44 · E-Mail: info@fluent.de

MESSEPRÄSENZ: **Simpat - Engine Expo u.a. – Wir freuen uns auf Ihren Besuch!**

2 Anforderungen

Dieser Beitrag konzentriert sich auf das parametrisierte Optimierungsproblem, das in der Literatur [5] auch Vollrastrabstimmung genannt wird: Für jeden Betriebspunkt sollen ein oder mehrere Stellgrößenkombinationen gefunden werden, die optimal für eine gegebene Zielfunktion sind, beispielsweise für den spezifischen Kraftstoffverbrauch. Dabei müssen mehrere motorspezifische Besonderheiten beachtet werden:

- Messungen am Prüfstand sind teuer und zeitaufwändig. Daher ist eine geringe Zahl von Messungen wünschenswert.
- Die Messergebnisse sind verrauscht. Messfehler müssen also behandelt werden.
- Das Verändern mancher Stellgrößen bedingt relativ lange Einschwingzeiten des Systems. Insbesondere Betriebspunktwechsel sollten daher möglichst selten stattfinden.
- Es treten Limits, also Randbedingungen für die Optimierung auf.
- Besonders interessant sind sichere Optima, also Optima mit einer großen Umgebung guter Zielfunktionswerte.

Aufgrund dieser Punkte wird der in diesem Bericht beschriebene neue Algorithmus als modellbasiertes Optimierungssystem entworfen. Das globale Modell ist auf dem gesamten Suchraum definiert, der von der Betriebsebene und den Stellgrößen aufgespannt wird. Dadurch wird die Zahl der Messungen gering gehalten, außerdem ermöglicht das Modell die Behandlung von Messfehlern. Zwei Modelltypen werden verwendet: linear parametrisierte und nichtlineare Regressionsmodelle. Bei linear parametrisierten Modellen ist die Modellausgabe eine Linearkombination aus (nichtlinearen) Basisfunktionen, zum Beispiel bilden Polynome maximal dritten Grades das oft verwendete kubische Modell. Dadurch hängt die Modellausgabe linear vom Vektor der freien Parameter ab. Für einen gegebenen Messdatensatz lässt sich der optimale Parametervektor leicht eindeutig bestimmen, wenn der quadratische Fehler minimiert wird.

Nichtlineare Modelle sind dagegen ebenfalls nichtlinear im Parametervektor. Dies hat eine deutlich höhere Flexibilität zur Folge. Dafür muss nun der optimale Parametervektor durch eine nichtlineare

Optimierung, das Training, gewonnen werden. Ferner ist der optimale Parametervektor nicht mehr eindeutig bestimmt, und das Modell ist leicht in der Lage, sich den Messfehlern anzupassen (Overfitting). Daher müssen geeignete Regularisierungstechniken eingesetzt werden (zum Beispiel [6]).

Eine spezielle Klasse nichtlinearer Modelle sind Feed-Forward Netze, die zu den Neuronalen Netzen gehören [7], die für den vorgestellten Algorithmus eingesetzt werden.

3 Aktives Lernen

Unter dem Gesichtspunkt der hohen Dimensionalität des Suchraums und der Kosten für jede einzelne Messung besteht das Kernproblem der Motoroptimierung in einer guten Versuchsplanung. Ihre Aufgabe ist es, möglichst wenig Messpunkte so zu platzieren, dass möglichst viel Information gewonnen wird. Da dieser Informationsgewinn meistens statistisch quantifiziert wird, spricht man auch von statistischer Versuchsplanung. Es existieren sehr viele verschiedenartige Versuchsplanungstechniken und noch mehr Litera-

3 Aktives Lernen

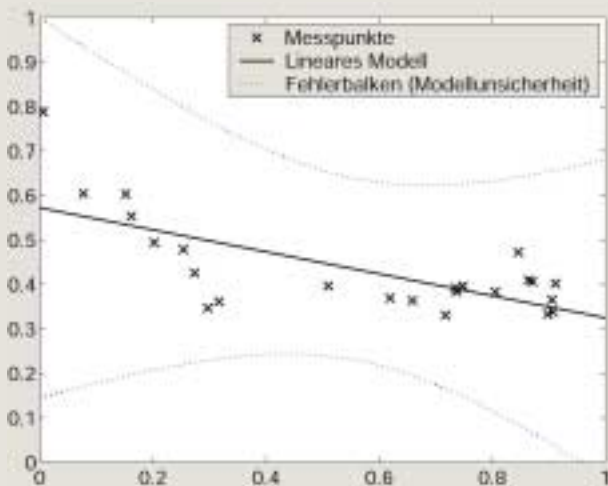


Bild 1: Lineare Modellierung eines Datensatzes mit Fehlerbalken. Die Modellunsicherheit wird am Rand des Suchraums maximal
 Figure 1: Linear approximation of a data set and error bars. The model uncertainty attains its maximum at the search space boundary

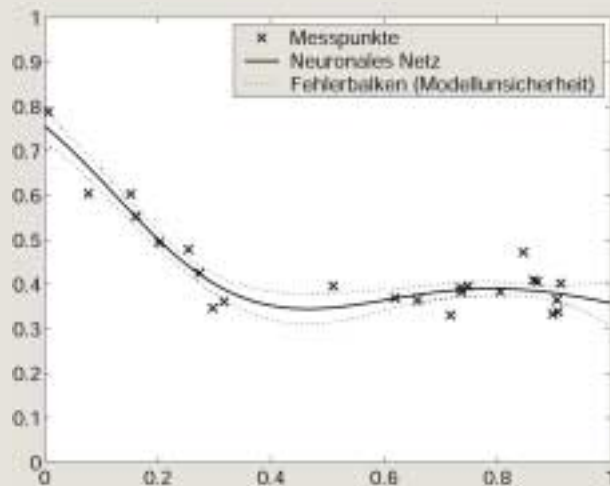


Bild 2: Modellierung eines Datensatzes mit einem neuronalen Netz (eine verdeckte Schicht mit drei hyperbolischen Neuronen, lineares Ausgabeneuron, und Shortcut-Verbindungen). Die Modellunsicherheit wird mit abnehmender Datendichte größer
 Figure 2: A data set is approximated by a Neural Network (on hidden layer with three hyperbolic neurons, linear output, shortcut connections). The model uncertainty grows with decreasing data density

tur zu dem Thema. Hier sei beispielhaft auf [8] verwiesen.

Eine wichtige Klasse von Methoden zur Versuchsplanung verwendet das Modell selbst. Dies geschieht, indem die Unsicherheit des Modells über den gesamten Suchraum ausgewertet wird. Bereits an dieser Stelle ergeben sich unterschiedliche Versuchsplanungskriterien: Beispielsweise kann man den Versuchsplan so anlegen, dass die durchschnittliche Modellunsicherheit möglichst klein wird, oder man minimiert die maximale Modellunsicherheit.

Die Auswertung der Modellunsicherheit an einem Punkt im Suchraum hängt vom Modell selbst ab. Zum Beispiel ist für ein lineares Modell die Unsicherheit immer am Rand des Suchraums am größten, **Bild 1**. Umgekehrt liegen in diesem Fall bei einem optimalen Versuchsplan die Messpunkte immer am Suchraumrand. Diese Überlegungen lassen sich auf linear parametrisierte, also beispielsweise kubische, Modelle ausweiten. Die Modellunsicherheit an einem Punkt im Suchraum kann dann durch einen Term ausgedrückt werden, der außerdem vom Modell und von den bereits gemessenen Punkten abhängt [8]. Ein wichtiges und leicht auswertbares Optimalitätskriterium für einen Versuchsplan ist die Forderung, dass die Determinante der so genannten Informationsmatrix maximal wird. Solche Versuchspläne werden auch D-optimal genannt, und unter geeigneten Voraussetzungen ist dies äquivalent zu der Bedingung, dass die maximale Modellunsicherheit über den Suchraum minimal ist [8].

Es fällt auf, dass in dieser Darstellung für die Modellunsicherheit an keiner Stelle die Messergebnisse vorkommen. Dies bedeutet, dass ein optimaler Versuchsplan bereits am Anfang erstellt werden kann und trotz Informationsgewinn durch die Messungen nicht weiter verbessert werden kann. Mit anderen Worten, diese Theorie favorisiert die Offline-Optimierung. Die Situation ändert sich jedoch grundlegend, wenn nichtlineare Modelle betrachtet werden. Auch in diesem Fall kann die Modellunsicherheit mit einer Informationsmatrix ausgedrückt werden, diese ist jedoch nun von dem Ergebnis des Trainings des Modells (beispielsweise eines Neuronalen Netzes) abhängig, also insbesondere auch von den Messergebnissen [9].

Werden nichtlineare Modelle für die Approximation der Zielfunktion verwendet, so kann also eine Online-Versuchsplanung die durch Messungen gewonnenen Informationen nutzen, um neue Messpunkte besser zu platzieren. Wir betrach-

ten daher jetzt die folgende Situation: Es liegen Messdaten vor, und gesucht ist ein weiterer Punkt im Suchraum, der derart bestimmt werden soll, dass der (erwartete) Informationsgewinn durch Messung maximal ist. Dieses Szenario wird auch als aktives Lernen (Active Learning) oder Query bezeichnet, da das Modell selbst entscheidet, welcher Punkt als nächstes trainiert wird. Ein mögliches Kriterium ist die Maximierung der Modellunsicherheit, die durch die Informationsmatrix und den

Punkt selbst gegeben ist. Dies hängt wiederum eng mit dem Kriterium der D-Optimalität zusammen. Ein Beispiel ist in **Bild 2** zu sehen: Die Unsicherheit des Modells (hier ein neuronales Netz) ist in Bereichen mit niedriger Datendichte größer.

Von diesem Kriterium gibt es noch einige Varianten [9, 10], die unter leicht unterschiedlichen theoretischen Annahmen gewonnen werden und die in der Praxis sehr ähnlich sind. Jedoch neigen diese Verfahren zu ungünstigem Verhalten, wenn

3 Aktives Lernen

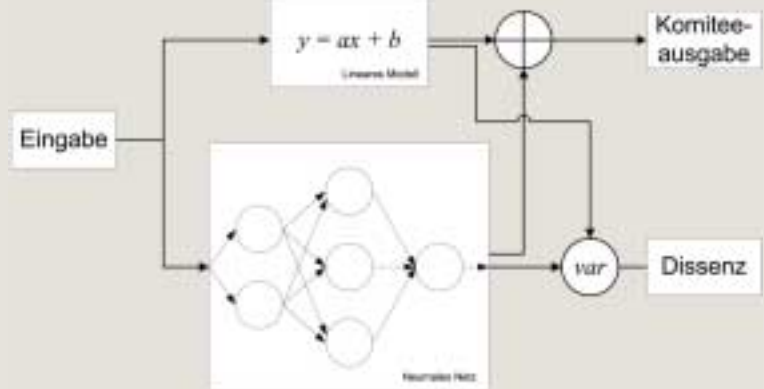


Bild 3: Schematische Darstellung eines Modellkomitees aus einem linearen Modell und einem neuronalen Netz
Figure 3: Sketch of a committee consisting of a linear model and a Neural Network

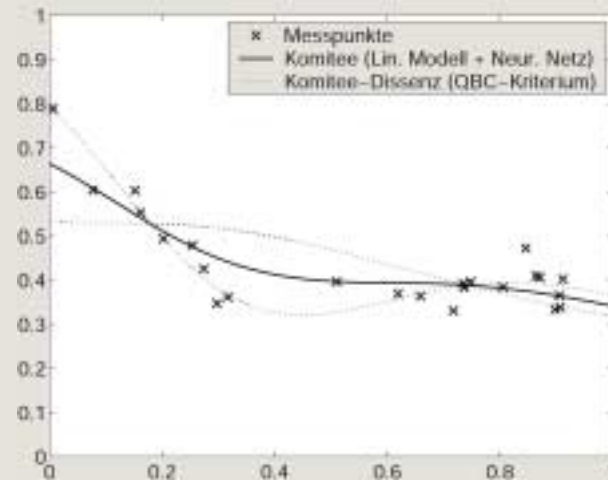


Bild 4: Modellierung eines Datensatzes durch ein Komitee aus linearer Regression und neuronalem Netz. Die Modellunsicherheit (QBC-Kriterium) ist dort maximal, wo die Modelle am wenigsten übereinstimmen (vgl. Bild 1 und Bild 2)

Figure 4: A committee of a linear model and a Neural Network approximates a data set. The model uncertainty (QBC criterion) is large in regions where the committee members do not agree (cf. Figure 1 and Figure 2)

die Annahme der Normalverteilung nicht erfüllt ist. Genauer wird die Bestimmung der Modellunsicherheit, wenn große Samples von Modellparametern entsprechend ihrer wahren Wahrscheinlichkeitsverteilung gebildet werden. Diese Markov-Chain-Monte-Carlo-Methoden (MCMC, siehe zum Beispiel [11]) sind jedoch sehr zeitaufwändig.

Eine weitere Möglichkeit besteht darin, ein Komitee von Modellen zu trainieren. Dies hat mehrere Vorteile: Erstens ist der erwartete Modellfehler im Komitee kleiner als für einzelne Modelle, zweitens können verschiedene Modelltypen gleichzeitig betrachtet werden (heterogene Komitees), und drittens kann die Nichtübereinstimmung oder Dissens (also die Varianz) der Ausgaben der Komiteemitglieder als Maß für die Modellunsicherheit, also als Query-Kriterium verwendet werden. Dieses Query-by-Committee-Kriterium (QBC) [12] war in unseren empirischen Vergleichen [13] den anderen Kriterien überlegen

und ist zusätzlich extrem schnell auszuwerten, daher wird es im neuen Algorithmus verwendet. Übrigens lässt sich QBC in homogenen Komitees auch als Variante einer MCMC-Methode auffassen: Jedes Komiteemitglied ist ein Sample des Parametervektors, wobei aufgrund des Trainingsprozesses die Wahrscheinlichkeitsverteilung zu ihren Maxima hin verzerrt ist. Die Verwendung von heterogenen Modellkomitees verringert ferner die Wahrscheinlichkeit, dass auf Grund eines ungeeigneten Modelltyps die Modellausgabe, obwohl einigermaßen mit den Trainingsdaten übereinstimmend, völlig falsch ist. **Bild 3** zeigt schematisch ein Komitee aus einem linearen Modell und einem neuronalen Netz. In **Bild 4** sieht man die Modellausgabe eines solchen Komitees und den Dissens der Komiteemitglieder, also das QBC-Kriterium.

Im Verlauf der Online-Optimierung sollte nicht nur die Modellunsicherheit, sondern auch in zunehmendem Maß das

bereits erworbene Wissen über die Zielgröße, also das aktuelle Modell, in das Query-Kriterium einfließen. Dadurch kann der Suchraum in Bereichen, in denen die Zielfunktionswerte gut sind, gründlicher exploriert werden, und die Optima werden genauer lokalisiert. Im vorgestellten Algorithmus geschieht dies, indem Teile des Suchraums ausgeschlossen werden, in denen die Zielfunktionswerte schlechter als ein Schwellenwert sind. Der mbminimize-Algorithmus wird mit einem initialen Versuchsplan gestartet, der nach raumfüllenden oder D-optimalen Kriterien berechnet wird. Dessen Ablauf wird mit den in [14] beschriebenen Methoden optimiert. Danach werden mittels aktiven Lernens weitere Messpunkte bestimmt, gegebenenfalls in ihrer Reihenfolge optimiert und vermessen. Schließlich werden die Modelloptima bestimmt und verifiziert.

4 Zusammenfassung

Der mbminimize-Algorithmus wurde gründlich an Hand von Testfunktionen und Simulationen ausgewertet und zeigt sehr gute Optimierungsergebnisse. Beispielsweise werden alle Optima einer veräuschten Branin-Funktion [15] mit weniger als 40 Funktionsauswertungen korrekt lokalisiert. Auch erste Versuche am Prüfstand ergaben vielversprechende Ergebnisse.

Durch die Verwendung eines globalen Modells und durch den Einsatz hochentwickelter Online-Versuchsplanungstechniken, insbesondere aktiven Lernens, hat der in diesem Beitrag vorgestellte mbminimize-Algorithmus das Potenzial, die Anzahl der benötigten Funktionsauswertungen gegenüber herkömmlichen Optimierungsverfahren deutlich zu reduzieren. Der Ansatz bietet ein breites Spektrum an Möglichkeiten der Erweiterung und Kombination mit anderen Techniken. Beispielsweise können andere Modelltypen integriert werden. Ein anderer Ansatzpunkt betrifft die Erweiterung um multi-kriterielle Techniken für die Optimierung von mehreren Zielgrößen. In Teil 2 dieses Beitrags wird die Behandlung von Suchraumgrenzen vorgestellt werden.

Diese Arbeit wurde vom BMBF gefördert (Kennziffer 01 IB 805 A/1).

Literaturhinweise

[1] Mitterer, A.: Optimierung vielparametrischer Systeme in der Antriebsentwicklung, statistische Versuchsplanung und künstliche neuronale Netze in der Steuergeräteauslegung zur Motorabstimmung. Dissertation, Lehrstuhl für Meßsystem- und Sensortechnik, TU München, 2000

MAXIMATOR®
Wir machen Druck!
4000 bar

Produktionsanlagen für das hydraulische Fügen von Nockenwellen

Dieses Verfahren hat gravierende Vorteile gegenüber geschmiedeten Teilen.

Durch den hohen hydraulischen Druck, der von innen auf die Welle wirkt, werden die vorgefertigten Nocken mit der Welle kraftschlüssig verbunden.

- geregelter Druckaufbau
- hohe Reproduzierbarkeit

MAXIMATOR®
Postfach 7 • D-37449 Zorge
Tel. 05586 / 803-0 • Fax 803-40
Internet: www.maximator.de
eMail: info@maximator.de

TECHNISCHE BÜROS

Nord	Tel. 02562 / 8168-0
Ost	Tel. 0341 / 6014227
Süd	Tel. 07261 / 9454-0
West	Tel. 02052 / 888-0

- [2] Schüler, M.; Hafner, M.; Isermann, R.: Einsatz schneller neuronaler Netze zur modellbasierten Optimierung von Verbrennungsmotoren. Teil 1: Modellbildung des Motor- und Abgasverhaltens. In: MTZ 61 (2000) Nr. 10, S. 704-710
- [3] Hafner, M.; Schüler, M.; Isermann, R.: Einsatz schneller neuronaler Netze zur modellbasierten Optimierung von Verbrennungsmotoren. Teil 2: Stationäre und dynamische Optimierung von Verbrauch und Emissionen. In: MTZ 61 (2000) Nr. 11, S. 2-8
- [4] Gschweidl, K.; Pfluegl, H.; Fortuna, T.; Leithgoeb, R.: Steigerung der Effizienz in der modellbasierten Motorenapplikation durch die neue CAMEO Online DoE-Toolbox. In: ATZ 103 (2001)
- [5] Bredenbeck, J.: Statistische Versuchsplanung für die Online-Optimierung von Verbrennungsmotoren. In: MTZ 60 (1999) Nr. 11, S. 740-744
- [6] MacKay, D. J. C.: A practical Bayesian framework for backpropagation networks. In: Neural Computation 4 (1992) S. 448-472
- [7] Zell, A.: Simulation neuronaler Netze. Addison-Wesley, Bonn, 1994
- [8] Bandemer, H.; Bellmann, A.: Statistische Versuchsplanung. Teubner, 4. Auflage 1994
- [9] Cohn, D. A.: Neural network exploration using optimal experiment design. In: Jack D. Cowan, Gerald Tesaro, and Joshua Alspector, (Editoren), Advances in Neural Information Processing Systems 6 (1994), S. 679-686
- [10] Paass, G.: Query sampling for prediction and model selection. In: IPSJ Magazine 38 (1997) Nr. 7, S. 562-568
- [11] Neal, R. M.: Probabilistic inference using markov chain monte carlo methods. Technical report, Dept. of Computer Science, University of Toronto, Sep 1993. CRG-TR-93-1
- [12] Seung, H.; Opper, M.; Sompolinski, H.: Query by committee. In: Fifth Annual Workshop on Computational Learning Theory, 1992 (ACM, New York), S. 287-294
- [13] Poland, J.; Zell, A.: Different criteria for active learning in neural networks: A comparative study. In: M. Verleysen (Editor), Proceedings of the 10th European Symposium on Artificial Neural Networks (2002), S. 119-124
- [14] Knödler, K.; Poland, J.; Mitterer, A.; Zell, A.: Optimizing data measurements at test beds using multi-step genetic algorithms. In: N. Mastorakis (Editor), Advances In Fuzzy Systems and Evolutionary Computation (2001), S. 277-282
- [15] Dixon, L. C. W.; Szego, G. P.: The optimization problem: An introduction. In: Towards Global Optimization II. North Holland, 1978

Morgen ist Heute!



www.all4engineers.com

Das Wissensportal für Automobil-Ingenieure

Bestellen Sie jetzt Ihr persönliches Newsletter-Abo – kostenlos!

Mit dem all4engineers-Newsletter entgeht Ihnen kein Branchen-Highlight: Top-Infos, Trends und Termine, die Ingenieure interessieren, motivieren und inspirieren. Jede Woche neu. Jede Woche Wissen pur. Holen Sie sich den Newsletter. Auf der Website: www.all4engineers.com oder direkt per E-Mail: all4engineers@bertelsmann.de.

Erfahren Sie heute, was die Automobiltechnik morgen bewegt: ■ Visionen ■ Aktuelle News zu Branchentrends ■ Suchmaschine für alle Inhalte von ATZ, MTZ, AutoTechnology und Automotive Engineering Partners ■ Jobbörse

all4engineers

ATZ

MTZ

AutoTechnology

ENGINEERING PARTNERS

For an English version of this article, see **MTZ worldwide**. For information on subscriptions, just call us or send an email or fax.



MTZ

Vieweg Verlag · Postfach 1546 · D-65173 Wiesbaden
Hotline 06 11778 78-151 · Fax 06 11778 78-423
email: vieweg.service@bertelsmann.de