Journal of Mechanical Design

Guest Editorial

Special Issue: Data-Driven Design (D^3)

Introduction

With the arrival of cyber-physical systems or "internet of things" era, massive human- and machine-generated data will create unprecedented challenges and at the same time unmatched opportunities in advancing the theory, methods, tools, and practice of data-driven design for products, systems, and services. By exploiting such huge, versatile, and highly contextualized through-life data, design engineers can harness their organization's competitive edge by uncovering patterns, novel insights, and knowledge for data-driven design. The aim of this special issue is to bring together original and archival articles that present significant contributions in advancing the field of data-driven design.

The initial idea of the special issue originated from a discussion among Ying Liu, Yan Wang, and Charlie Wang during ASME IDETC/CIE 2016 in Charlotte and was enthusiastically supported by *Journal of Mechanical Design*'s Editor, Dr. Shapour Azarm, and Harrison Kim, who later was invited to join as one of the guest editors. Through a world-wide dissemination of the special issue's call for papers, we received 85 submissions, among which 36 papers were selected for peer review evaluation by a minimum of three reviewers. After a minimum of two rounds of review, 20 papers of different types were accepted for publication, including 18 research papers, one review paper, and one technical brief. The key topics among the accepted papers are: D^3 methods (foundation and principles), variability and uncertainty in D^3, team dynamic in D^3, D^3 and lifecycle, and D^3 applications and case studies with overarching utilization of data from a wide variety of sources and with different magnitude and sizes. The level of enthusiastic response to our call for papers from the design engineering community confirmed our belief that the community is poised to take a leadership role in advancing knowledge and application domains of D^3. Below is provided a short summary of the papers in this issue following the previously mentioned subgrouping of the topics.

D^3 Methods

The paper, "An Integrated Approach for Design Improvement Based on Analysis of Time-Dependent Product Usage Data" by Ma et al., illustrates that product usage data are used to monitor the performance of functional modules in a hierarchical view. The kernel principal component analysis is applied to reduce the dimension of time-dependent performance feature data set corresponding to functional modules, before Gaussian mixed model fitting is used to model degradation under uncertainty. Health degradation severity of each function thus can be characterized from the distributions. To-be-modified design parameters are then identified from the functions with severe degradation tendencies.

The paper, "A Data-Driven Text Mining and Semantic Network Analysis for Design Information Retrieval," by Feng Shi et al., proposes an approach of ontology-based design concept "wordnet" to address some of the current limitations in design document retrieval. The key technique relies on text mining to establish an unsupervised learning ontology network. Validation through an engineering design case study shows that the proposed approach is able to recognize those highly related complex design tasks and their associations with different engineering elements.

The paper, "V4PCS: Volumetric 4PCS Algorithm for Global Registration," by Huang et al., provides a geometry registration algorithm that helps identify the similarity between two designs, where the optimum alignment of two surface tessellation models is found efficiently from sampled vertices. The paper shows that the design comparison can support rapid product customization.

The paper, "A Systematic Function Recommendation Process for Data-Driven Product and Service Design," by Zhang et al. presents a systematic function recommendation process to suggest new functions to an existing product and service. Different from the conventional approaches where new functions are largely formulated by experienced designers, the proposed approach builds upon recommendation systems that dynamically catch the trendy requirements from targeted users that are not recognized by existing product and service yet. A detailed case study reveals the merits of the proposed approach.

The paper, "Beyond the Known: Detecting Novel Feasible Domains Over an Unbounded Design Space," by Chen and Fuge, presents a data-driven adaptive sampling technique—ε-margin sampling—to discover feasible domain in an unbounded design space in an efficient manner. The method both learns the domain boundary of feasible designs, while also expanding the knowledge of the design space as available budget increases. The authors also couple design manifolds with ε-margin sampling to actively expand high-dimensional design spaces without incurring the exponential penalty. The approach is demonstrated in real-world examples of glassware and bottle design cases.

The paper, "Data-Driven Sizing Specification Utilizing Consumer Text Reviews," by Chaklader and Parkinson, introduces a new method to determine preliminary design specifications related to human–artifact interaction. The proposed new approach primarily uses text mining of a large number of consumer reviews to suggest human variability information that is essential for interaction. A weighted phrase rating metric is studied which does not require any human intervention but quickly and economically provides information useful to the establishment of design specifications.

The paper, "Automated Extraction of Function Knowledge From Text," by Cheong et al., develops a method to automatically extract function knowledge from natural language text. The extraction method uses syntactic rules to extract subject–verb–object triplets from parsed text. Then, the Functional Basis taxonomy, WordNet, and word2vec were leveraged to classify the triplets as artifact-function-energy flow knowledge. The method can find function definitions for 66% of the test artifacts. For those artifacts found, 50% of the function definitions identified are compiled in a well-known design repository. In addition, 75% of the

most frequent function definitions found by the method are also defined in the same design repository.

The paper, "A Convolutional Neural Network Model for Predicting a Product's Function, Given Its Form," by Dering and Tucker, introduces a deep learning approach based on three-Dimensional Convolutions that predicts functional quantities of digital design concepts. Case studies have been presented in this paper to verify the research questions that are derived from this work, including whether the learned 3D convolutions are able to accurately calculate the functional quantities, determine what the latent features discovered by this network mean, and assess whether the proposed model can perform better than other deep learning approaches.

The paper, "Mitigating Online Product Rating Biases Through the Discovery of Optimistic, Pessimistic, and Realistic Reviewers," by Lim and Tucker, offers a new method to lower user rating biases that are caused by customers' optimism or pessimism. By considering the rating history and tendency of a reviewer, the work backed by an unsupervised model aims to adjust the influence on ratings in order to provide customers a more objective and accurate feedback.

Variability/Uncertainty in D^3

The paper, "Modeling the Variability of Glenoid Geometry in Intact and Osteoarthritic Shoulders," by de Vries and Parkinson, presents a research work to model the geometric variability of the glenoid of the scapula. The pipeline based on geometric fitting, radial basis functions, and principal component analysis, which can represent the glenoid in a new manner. The work was validated against existing approaches and CT scans from 42 patients. The models created is expected to help surgeons and engineers to understand the effects of osteoarthritis on bone geometry, as well as the range of variability present in healthy shoulders.

The paper, "A Taylor Expansion Approach for Computing Structural Performance Variation From Population-Based Shape Data," by Wang and Qian, investigates a Taylor expansion based method for computing structural performance variation over its shape population. To overcome the potential inaccuracy of Taylor expansion for highly nonlinear problems, a multipoint Taylor expansion technique is proposed in the paper, where the parameter space is partitioned into different regions and multiple Taylor expansions are locally conducted. It works especially well when combined with the dimensional reduction of the principal component analysis in the statistical shape modeling.

The paper, "Mining Process Heuristics from Designer Action Data via Hidden Markov Models," by McComb et al., shows an application of data-mining techniques to quantitatively study the processes that designers use to solve configuration design problems that are characterized by the assembly of components into a final desired solution. The extraction of human problem-solving heuristics is automated through the application of hidden Markov models, which show that designers proceed through four procedural states in solving configuration design problems.

The paper, "Predicting Future Importance of Product Features Based on Online Customer Reviews," by Jiang et al., illustrates that opinion mining is adopted to extract product features from customers' reviews. Fuzzy sets and rules are used to accommodate the imprecision of natural languages. The importance levels or weights of different product features are determined through fuzzified frequencies and sentiment scores. The fuzzy time series method is also applied to predict future importance weights.

Team Dynamics in D^3

The paper, "Concept Clustering in Design Teams: A Comparison of Human and Machine Clustering," by Zhang et al., presents a machine learning tool to cluster design concepts and compares the outcome to that of manual clustering. The goal of the clustering algorithm is to support design teams in identifying possible areas of "over-clustering" and/or "under-clustering" in order to

enhance divergent concept generation process. The approach was demonstrated by the data generated in a graduate new product development class.

D^3 and Lifecycle

The paper, "Visual Analytics Tools for Sustainable Lifecycle Design: Current Status, Challenges, and Future Opportunities," by Ramanujan et al., provides a review of previous research that has created visual analytics tools in sustainable lifecycle design and highlights existing challenges and future opportunities. The opportunities are highlighted for different stages of lifecycle—design, manufacturing, distribution and supply chain, use-phase, and end-of-life.

The paper, "InnoGPS for Data-Driven Exploration of Design Opportunities and Directions: The Case of Google Driverless Car Project," by Luo et al., demonstrates that patent mining techniques can be applied to identify technological neighborhoods by analyzing proximity of patent domains in graph models. Future design and technological opportunities can be discovered by adopting the proposed method.

D^3 Applications and Case Studies

The paper, "Data-Driven Styling: Augmenting Intuition in the Product Design Process Using Holistic Styling Analysis," by Ranscombe et al., proposes the Holistic Styling Analysis (HSA) for improved digital shape comparison applied to 3D geometry of products. HSA provides objective assessment of difference in appearance to form the basis for designers to rationalize styling to other stakeholders during the design process. The approach enables styling designers to use data to drive their activities in the same manner as other stakeholders. An automotive case study validates the proposed approach by providing objective reference measures for differentiation in multiple products.

The paper, "Identification of Performance Requirements for Design of Smartphones Based on Analysis of the Collected Operating Data," by Zhang et al., showcases the case when designers of smartphones analyze the operating data for CPU performance and utilization. A sigmoid like function is used to approximate the cumulative distribution function in order to identify customer satisfaction and the point of cost effectiveness.

The paper, "Dynamic Data-Driven Design of Lean Premixed Combustors for Thermoacoustically Stable Operations," by Chattopadhyay et al., uses collected experimental data to generate stability map of combustor in design parameter space. Support vector machine and Markov models are used to identify system states. The relationship between operational conditions and system response is built with a Gaussian process regression. Designers can then use such relationship to perform design optimization.

The paper, "Mining Patent Precedents for Data-driven Design: The Case of Spherical Rolling Robots," by Song and Luo proposes a heuristic approach to patent data-driven design and demonstrates the approach in a case study of spherical rolling robot. The approach is an iterative and heuristic methodology to exhaustively search for patents as precedents of the design of a specific technology or product for data-driven next design. Designers can utilize the methodology to make sense of retrieved patent data to explore new design opportunities.

Finally, the guest editorial team would like to take this opportunity to thank all of the contributing authors for their excellent work. We are also very grateful to the reviewers for offering their precious time and efforts and for providing constructive comments in a timely manner, especially for those submissions that were reviewed for three or even four rounds. Without all of you, this special issue would not be possible. Enjoy reading the papers!

Harrison Hyung Min Kim
Department of Industrial and
Enterprise Systems Engineering,
University of Illinois
at Urbana-Champaign,
Urbana, IL 61801

Ying Liu
Institute of Mechanical
and Manufacturing,
School of Engineering,
Cardiff University,
Cardiff CF24 3AA, UK
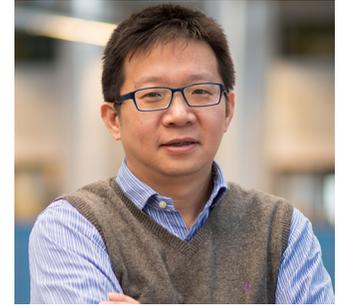
Charlie C. L. Wang
Department of Design
Engineering,
Delft University of Technology,
Delft 2628CE, The Netherlands

Yan Wang
George W. Woodruff School of
Mechanical Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332