Ian Tseng
e-mail: iht@alumni.cmu.edu

Jonathan Cagan
e-mail: cagan@cmu.edu

Department of Mechanical Engineering,
Carnegie Mellon University,
5000 Forbes Ave.,
Pittsburgh, PA 15213

Kenneth Kotovsky
Department of Psychology,
Carnegie Mellon University,
Baker Hall 342c,
Pittsburg, PA 15213
e-mail: kotovsky@cmu.edu

Concurrent Optimization of Computationally Learned Stylistic Form and Functional Goals

Great design often results from intelligently balancing tradeoffs and leveraging of synergies between multiple product goals. While the engineering design community has numerous tools for managing the interface between functional goals in products, there are currently no formalized methods to concurrently optimize stylistic form and functional requirements. This research develops a method to coordinate seemingly disparate but highly related goals of stylistic form and functional constraints in computational design. An artificial neural network (ANN) based machine learning system was developed to model surveyed consumer judgments of stylistic form quantitatively. Coupling this quantitative model of stylistic form with a genetic algorithm (GA) enables computers to concurrently account for multiple objectives in the domains of stylistic form and more traditional functional performance evaluation within the same quantitative framework. This coupling then opens the door for computers to automatically generate products that not only work well but also convey desired styles to consumers. [DOI: 10.1115/1.4007304]

1 Introduction

Since their invention, computers have gained many new capabilities. Originally, capable only of simple logic and arithmetic, computers were best suited for repetitive and tedious calculation. Modern software technology has now expanded this capability to enable computers to optimize design parameters and generate new functional designs. One key limitation holding back complete computational design generation of products, is a computer's inability to understand human aesthetic judgments. When calculating easily quantifiable measures, such as the coefficient of drag of a particular shape, computers are efficient, effective, and can often outperform humans. On the other hand, computers have been unable to judge many less quantifiable measures such as the beauty of a piece of artwork or the apparent sportiness of a car. These seemingly trivial judgments for humans are not understandable by computers since computers can only process objective representations, and not subjective representations.

Finding a way for computers to understand and quantify these subjective human perceptive judgments opens the door for computers to design not only the functional elements of a product, but to generate and optimize the form elements as well. The work described in this paper aims to bridge the gap between computational processing and stylistic form with a framework that uses artificial neural networks (ANN) to enable computers to learn consumer stylistic form desires from survey data. Then including both this stylistic form judgment via the ANN as a fitness function, and more traditional functional evaluation, a genetic algorithm is used to generate new designs that satisfy consumer desires for both functional performance and stylistic desire. The methods in this research can be applied to the design of a wide variety of product categories. Nevertheless, in order to elicit responses from consumers and to better illustrate the abilities of the methods, an illustrative design problem is needed. An appropriate design problem should be one with a richly interactive relationship between stylistic form and function, and should also be applicable and familiar to all participants. A vehicle styling design problem is used as an illustrative platform for testing the relationship between stylistic form and function. Stylistic interpretation in this work includes not only aesthetics (beauty) but also people's perception of ruggedness, fuel efficiency and sportiness.

In order to demonstrate this system's capabilities in concurrently optimizing both stylistic form goals and functional goals, a Pareto Frontier is mapped between two vehicle design goals, and vehicle designs are generated that target multiple combinations of stylistic and functional goals.

2 Background

The inherent disconnect between computational modeling and stylistic goals arise partially due to differences in representation. Most commonly, form is represented graphically and functional goals are represented numerically. A first step toward incorporating both goal types in a single analysis is a method for generating shape in a controlled deterministic manner that connects the shape being generated with a unique numerical representation, which enables direct and objective comparisons with functional goals, as well as the ability for shape representation and computer algorithms to interface. This desire to distill artistic form into absolute and mathematical measures and equations has motivated researchers for many decades [1]. This bridging of representations would enable direct and objective comparisons between aesthetic form and functional goals, which could pave the way for formalizing discussions of the relationship between form and function [2]. There are many methods that have been used successfully to accomplish this first step goal of representing shape deterministically, including shape grammars [3–8], morphing [9,10], parametric models [11], and visual decomposition [12]. Ranscombe et al. [12] used visual decomposition to represent vehicle designs, and varied the level of detail shown to participants to determine which details were most necessary for consumer judgment of vehicle brand, vehicle segment, and vehicle character. They concluded that additional details beyond the outline of the vehicle did not add to participants' ability to correctly judge the vehicle design, and that the outline of the vehicle design alone was sufficient to represent it. Furthermore, the authors concluded

Contributed by the Design Automation Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received August 11, 2011; final manuscript received June 19, 2012; published online October 2, 2012. Assoc. Editor: Bernard Yannou.

by highlighting a need for numeric metrics that can measure aesthetic features, which could provide guidance to designers, which is the one of the goals of the research in this paper.

Even with the aforementioned computational representations of form, computers are unable to understand what is required for the appearance of a design to convey style to consumers. In measuring and modeling consumer preference of functional characteristics of a design, much work has been done using decision-based design frameworks [13] such as conjoint analysis using logit models [14–16]. These methods use logistical regression to extract individual utility functions for each of the characteristics. These methods assume that none of the characteristics that are being modeled interfere with one another; in other words, the utility of the design to the consumer is simply the summation of the individual utility functions for each of the characteristics. When modeling features and characteristics that are independent, this method has been shown to approximate human preference well [7,17]. Unfortunately these methods are ill prepared to model interaction effects between the characteristics, as would be necessary in modeling stylistic desires of consumers.

Up until this point, the literature has focused on the blanket issue of preference, but when judging stylistic desires of consumers this idea of preference must be decomposed into separate stylistic judgments. Satisfying these stylistic judgment desires is an example of emotional design, which has been shown to affect consumer perceptions of quality and emotional experience for everyday objects, which in turn can affect their buying decisions [18]. Research has highlighted consumer sensitivities to the emotion of products being designed, for instance clothes irons that were designed to appear happy, cute, or tough, were perceived accordingly in studies [19]. A method that can be used to decompose product goals and emotions is Kansei engineering [20,21]. These required physical traits are often measured using surveys that employ the semantic differential. The semantic differential [22] is a way to measure the meaning of concepts and break them down into multiple dimensions of opposing adjectives using factor analysis. Due to the semantic differential's capability with abstract concepts, a common use is on surveys where consumers are asked to rate a product on a linear scale that ranges between two opposing words. The semantic differential has been widely used to survey emotional, aesthetic, and stylistic characteristics of the shape of designs, including the perceived environmental friendliness of vehicle silhouettes [23], the sportiness of vehicle designs [10], and generic shapes [24].

A method for generating form to satisfy desired stylistic goals was developed [25], which took a unique approach toward the requirements of shape generation by using a Fourier series to represent the genetics of car silhouettes. The car designs could then be refined to better fulfill user desires using an interactive genetic algorithm (IGA). A genetic algorithm (GA) is a genetically inspired optimization technique that creates a random population of designs, and refines the population for improved performance by calculating the fitness of each design to probabilistically remove weaker design attributes and reinforce stronger design attributes. IGAs, also called interactive evolutionary computation, are genetic algorithms that specifically employ the use of a human to perform the fitness judgments. IGAs allow people with little or no CAD experience to interact with software to create and modify designs and shapes. IGAs are particularly well suited in the generation of stylistic form and have been used successfully in the fields of fashion design [26], and in the shape design of eyeglass frames [27]. A survey paper that covers many applications and concerns for IGAs can be found in Ref. [28]. A major limitation of IGAs is that a human is required to perform the fitness evaluation, which violates two large goals of computational design: to reduce human involvement and to speed up processing time.

Thanks to advances in artificial intelligence, it is possible to train a computer to closely mimic many human responses. An ANN is a computational model that contains a network of interconnected artificial neurons or nodes arranged in layers [29]. This method was inspired by, and is loosely based on, how biological neural networks function. ANNs have been shown to have excellent abilities to recognize patterns in large data sets [30] and have been used successfully to perform character recognition [31]. ANNs have been coupled with parametric models and the semantic differential method to model human preference in office chair design [11] and bridge design [32]. ANNs are also able to circumvent the challenges of logit models when modeling interaction effects [29] because they are natively sensitive to interactivity between variables, as opposed to logit models, wherein the interactivity terms must be determined and set up separately. An example of ANNs modeling interactivity between variables is provided in this paper at the end of Sec. 5.2.

One limitation when trying to generate new designs using ANNs is their unidirectionality, meaning that while the neural network can mimic the decisions in the training data, it is impossible to reverse the direction and decode the input parameters from a desired output state. This differs from the regression methods described previously, which can be used directly to generate new designs. What is needed to complete the picture with ANNs is a method for network inversion. This desired network inversion can be achieved by coupling an artificial neural network with a genetic algorithm. An ANN can be used as the fitness function of the GA, which negates the need for the human fitness judgments as used in IGAs. This results in improved autonomy and speed, allowing new designs to be automatically bred in the GA that satisfy a desired output of the ANN. This method was used in Ref. [33] to evolve and generate computer images that were pleasing to users. This method has also been used with limited success on generating musically pleasing phrases in jazz music [34,35]. Researchers [36] further refined the technique into a model-based evolutionary computation system that uses a generic parametric model form, and is iteratively retrained using the explicit fitness function after a scheduled number of generations to retrain the neural network for greater accuracy around the optimal solution. It is assumed in this system that the fitness function is costly, meaning that it is either time or resource intensive to evaluate, so minimizing how often it needs to be evaluated is of large benefit. This is shown to be much more efficient and effective than other methods that retrain on other schedules, or that do not retrain at all.

Coupling ANNs and GAs with survey data built from five-level surveys to learn aesthetic form, much like the work presented in this paper, has been attempted by Tsutsumi and Sasaki [37] who paired this method with a five-level Kansei evaluation survey to design a gymnasium roof structure that satisfied both goals of aesthetic beauty and calculations of maximum stress. The work in this paper builds on the methods used and developed by Baluja et al. [33], Bull [36], and Tsutsumi and Sasaki [37], which use an ANN to capture people's aesthetic preference. In turn, this ANN then becomes the fitness function for the GA, which generates new designs that satisfy targeted consumer desires of stylistic form, and manages tradeoffs between functional performance and stylistic form. The resulting model can then be used in optimization with other objectives, such as functional characteristics, which can be traditionally represented and concurrently optimized using the GA.

An abstract framework similar to the one used in this paper is laid out in Ref. [38], which discusses the requirement of both an analysis algorithm and a synthesis algorithm in a design algorithm, whose goal is to produce an object that satisfies stylistic or artistic goals. In this framework, the purpose of an analysis algorithm is to determine how an object of a given description is interpreted and evaluated in terms of the interpretive conventions and criteria of the aesthetic system, whereas the purpose of a synthesis algorithm is to find the description of an object that fulfills some target goal of the interpretive conventions and criteria of the aesthetic system. In the research documented in this paper, an ANN acts as the analysis algorithm to evaluate designs based on survey results, while a GA acts as the synthesis algorithm, which searches for optimal designs using the evaluations of the analysis algorithm. The goal of this framework is to generate new designs that target specific desired aesthetic judgments.

3 Learning Stylistic Form From Survey Data

A two-stage experiment was conducted where participants were asked to rate computer generated car profiles for sportiness, ruggedness, beauty, and fuel efficiency. The first three judgment criteria, sportiness, ruggedness, and beauty, were selected as common subjective terms to describe the styling of vehicles. The fourth term, fuel efficiency, was chosen to assess whether a specific vehicle design aesthetic is associated with judgments of fuel efficiency, and whether a computer can learn this functional perception. Of note, in this paper we include not only aesthetics, but also people's interpretation of a vehicle's performance, based only on how it looks, which we call its stylistic form. In the first stage, participants generated survey data, which was used to train four ANNs per participant, one for each of the four rating categories. The resulting ANNs were then inverted using a GA in order to generate new designs that the ANNs rate highest and lowest in sportiness, ruggedness, beauty, and fuel efficiency. In the second stage, the participants were surveyed with the cars generated using data from their ratings to verify the performance of the ANN and GA learning and generation system. For more detailed information about this study, please refer to Tseng [39].

3.1 Participants. Eighteen volunteers participated in this experiment ranging from 22 to 34 years of age, and a split of 11 males and 7 females.

3.2 Procedure. For both stages of the experiment, participants were presented with a computer interface written in MATLAB as shown in Fig. 1. The interface displays a vehicle design generated using a parametric model and asks participants to rank whether each of four words presented, Sporty, Rugged, Beautiful, and Fuel Efficient, described the vehicle design Very Poorly, Poorly, Neither, Well, or Very Well. Participants of this study were first presented with a preview of 20 cars that were randomly generated using the parametric model in order for participants to calibrate their judgments and familiarize themselves with the task.

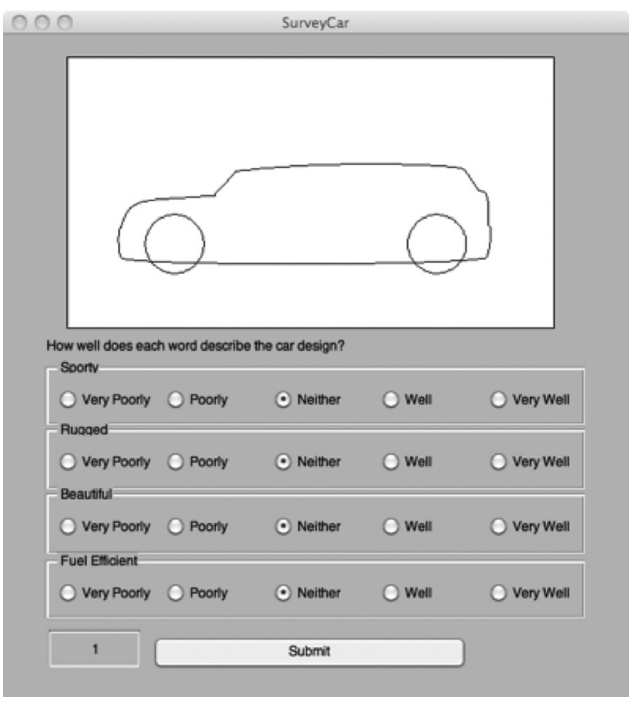

Fig. 1 Sample MATLAB survey

Based on performance during pilot studies, participants were also encouraged to form a personal set of mental preference rules that govern the reasoning behind their judgments, and were encouraged to stick with their rules through both stages of the experiment to maximize the consistency of their judgments. These mental rules were purely to aid the participants in making meaningful judgments. They were not shared with the researchers, and did not factor into the data collection or analysis.

The first stage survey presented participants with 276 vehicle designs that were used to train the neural networks. This first stage survey took an average of approximately an hour for participants to complete. Participants were encouraged to take breaks whenever they felt fatigued, with at least one break in the middle recommended. In between the two stages, the learning and generation system was trained on the data from the first stage of the experiment and verification designs were generated. After generating the verification designs, participants were asked to return for the second stage where they were presented with the 50 generated verification designs to assess the performance of the learning and generation system. This second stage verification survey took an average of approximately 15 min.

3.3 Design and Materials. For the first stage, each participant was presented with 276 computer generated vehicle designs to rate on a five-level scale in categories of sportiness, ruggedness, beauty, and fuel efficiency. The 276 designs presented consisted of 256 vehicles that were randomly generated in pre-assigned parameter ranges as discussed in Sec. 3.3.2, and 20 consistency verification car designs as discussed in Sec. 3.3.3. Details about the pre-assigned parameter range vehicle designs are discussed later. The order of the 256 vehicle designs with pre-assigned parameter ranges was randomized, and then the 20 consistency verification designs were inserted uniformly throughout the survey. The types of vehicles shown in the first stage of the experiment are shown in Table 1.

For the second stage, each participant was presented with a survey similar to the first stage, but with 50 computer-generated vehicle designs. 24 of the 50 designs presented consisted of vehicles designed by the learning and generation system to elicit high and low ratings in each of the four categories as discussed in Sec. 3.3.4, and 12 designs were the same consistency verification designs that were presented in the first stage, as discussed in Sec. 3.3.3. The remaining 14 designs were randomly generated and were only presented to help disguise the 24 generated test designs and the 12 consistency verification designs. The order of the vehicles presented were randomized, and any identical designs that appeared less than five questions apart were manually moved until no such violations were found. The types of vehicles shown in the second stage of the experiment are shown in Table 2.

3.3.1 Parametric Design Generation. In order to consistently and deterministically generate a large variety of car designs, a parametric car model was created. This model consists of eight cubic Bezier curves and two circles. Curve 1 defines the front bumper, Curve 2 defines the front grille and hood, Curve 3 defines the front windshield, Curve 4 defines the roof, Curve 5 defines the rear windshield, Curve 6 defines the trunk lid, Curve 7 defines the rear bumper, and Curve 8 defines the under-pan of the car, as labeled in Fig. 2. These eight curves are defined by twelve design

Table 1 Vehicles shown in experiment first stage (shown in random order)

	Number shown	Discussed in sections
Randomly generated vehicles in pre-assigned parameter ranges	256	3.3.2
Consistency verification designs	20	3.3.3
Total	276	—

Table 2 Vehicles shown in experiment second stage (shown in random order)

	Number shown	Discussed in sections
Model verification vehicle designs	24	3.3.4
Consistency verification designs	12	3.3.3
Random vehicle designs	14	—
Total	**50**	—

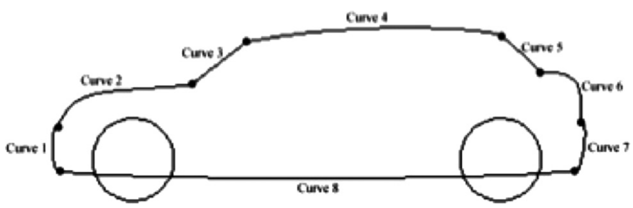

Fig. 2 Curves in vehicle design model

Table 3 Car design parameters

1	Belt angle – The angle of rise of the belt line from nose to tail.
2	Nose angle – The angle of rake of the nose of the car.
3	Ground clearance – The distance from the floor pan to the ground.
4	Body height – The height where the hood meets the windshield.
5	Roof height – The height of the top of the windshield.
6	Hood length – The length from the back of the bumper to the windshield.
7	Trunk length – The length from the back of the rear windshield to the front of the rear bumper.
8	Windshield rake angle – The angle the windshield leans back.
9	Rear windshield rake angle – The angle the rear windshield leans forward.
10	Wheel size – The diameter of the wheels.
11	Front wheel position – The length distance of the front wheel from the origin.
12	Rear wheel position – The length distance of the rear wheel from the origin.

parameters that can be varied in integer steps between 0 and 99. The range of each design parameter was calibrated to extremes as historically seen in car designs, and are constrained to ensure stability of the model such that only geometrically feasible models can be generated. The model's stability is further ensured by extensive testing around the extremes of all parameter values. The twelve design parameters and their definitions are shown in Table 3. For more information about this parametric design model, please refer to Tseng [39].

3.3.2 Random Vehicle Designs With Pre-assigned Parameter Ranges. In order to ensure maximum diversity of body shape for the cars generated in the first stage of the experiment, parameter ranges were pre-assigned for 256 of the designs. The car design parametric model, as discussed earlier, takes values between 0 and 99 for each of the twelve parameters. Parameters 1, 2, 4, 5, 6, 7, 8, and 9 in the model pertain to vehicle body shape, while Parameters 3, 10, 11, and 12 pertain to ground clearance, front and rear wheel position, and wheel size. Each parameter relevant to vehicle body shape was assigned a random value in either the upper half of the range (50–100), or the lower half of the range (0–50). To ensure that the population of cars presented in the survey had a maximum diversity of body shape, all upper and lower combinations of parameter values for parameters that pertain to vehicle body shape were presented in the survey, resulting in 256 car designs (two levels, eight parameters $= 2^8$ designs). The remaining four parameters do not pertain to or affect body shape, and were randomly generated in the full 0–100 range for these 256 cases.

3.3.3 Consistency Verification Vehicle Designs. In order to test for participant consistency, four predetermined car designs were presented five times each to each participant in the first stage, and three times each to each participant in the second stage, spread out uniformly throughout the surveys. The predetermined car designs were an average sedan, made up entirely of centered parameter values of 50, and the three vehicles designed to resemble a 2009 Chevrolet Corvette, a 2006 Toyota Prius, and a 2010 Range Rover Sport. This consistency verification data can help to determine how consistent a participant's ratings are with his- or her-self when the same car is presented multiple times, how consistent participants are between stages, and how much participants varied between each other. This verification data are used later to filter the results of the study. It is assumed that those participants who were less able to consistently judge the same vehicle design presented multiple times have a more weakly defined opinion about that particular judgment category, and thus it would be difficult to reliably generate vehicle designs that satisfy them in that particular judgment category.

3.3.4 Model Verification Vehicle Designs. To test the performance of the learning and generation system, designs were generated to target each participant's highest and lowest ratings in the four categories of sportiness, ruggedness, beauty, and fuel efficiency. This results in a total of eight vehicle designs. Tailor-made surveys were created for each participant using the high and low rating vehicles created using his or her own ANN. Each vehicle design was presented to the participant three times during the course of the second stage verification survey, resulting in 24 of the 50 designs presented. The remaining 26 of the 50 designs were randomly generated and were only presented to help disguise the 24 generated test designs. The order was randomized, and then filtered to ensure that no two identical cars could be presented within five cars of each other.

3.4 Data Analysis. This paper does not seek to advance the art of ANNs, so a readily available off-the-shelf implementation was used. The survey data in the first stage of this experiment was used to train an ANN generated using the Neural Network Fitting Tool in MATLAB's Neural Network Toolbox. This tool creates a feed-forward network with one hidden layer with a variable number of hidden neurons. The network is trained using the Levenberg-Marquardt backpropagation algorithm, which is a numerical minimization algorithm for nonlinear functions and operates by interpolating between the Gauss-Newton and Steepest Gradient techniques. More information about this algorithm can be found in Refs. [40,41].

Because this research does not intend to advance the art of ANNs, a brute force method was employed to find the optimal network structure and to deal with the stochasticity of training. A search was conducted to find the optimal number of hidden neurons. For each rating category for each participant, eight networks were generated; the eight networks contained either 2, 3, 4, 5, 6, 10, 12, or 20 hidden neurons. The network was trained with a randomly selected 194 (70%) of the 276 samples, and then validated using 41 (15%) samples, and then tested using 41 (15%) samples. The network structure that scored the lowest error for the 41 test samples was assumed to have the optimal number of layers. The optimal number of layers varied from 3 hidden neurons to 20 hidden neurons with no pattern that was discernable to the researcher. To deal with system stochasticity, the network was trained and saved five more times, and the network with the lowest mean squared error for the test data sample was selected to generate verification car designs for the second stage verification survey. Four ANNs were created and trained for each participant, one for each category of judgment.

Once the ANNs were created, the GA tool in MATLAB's Optimization Toolbox was used to find vehicle designs that the ANNs rate as being the highest and lowest in each of the four categories. The GA optimization tool is designed to find the minimum of the fitness function within a set of constraints. The fitness functions used in each case were the four ANNs per participant created to model the survey data from the first stage. In order to find the lowest rated design, the output of the ANN was used as the fitness function, and in order to find the highest rated designs, the negative of the output of the ANN was used as the fitness function. Each chromosome being bred by the GA is made up by twelve genes, which correspond to the twelve vehicle design parameters. It was found by experimentation that a population of 20 chromosomes was sufficient to ensure convergence in this system. The 20 chromosomes for the initial population of designs used in the GA were randomly generated by MATLAB, and were bred using the default combination of selection, reproduction, mutation, crossover, and migration functions. In order to maximize the chance that the global optimum was found, the default stopping criteria was disabled, and the algorithm was allowed to run for 5000 iterations for each design, which was found by experimentation to be ample. Furthermore, the plot of each objective function was inspected by the experimenter to ensure convergence. To illustrate, the resulting highest and lowest rated designs in each of the four categories for one of the participants (Participant 18) are presented in Figs. 3–6.

As described in Sec. 3.3.4, the designs that were generated using the ANN and GA system to target high and low ratings were assembled into the verification surveys presented in the second stage survey. The generated verification designs were shown three times each to eighteen participants, resulting in 108 total ratings for each category, 54 for highly rated designs, and 54 for low rated designs. These ratings for specifically targeted designs were compared against the average ratings for all survey responses from the first stage survey, for each specific category. These average survey responses contain 276 responses for each participant, or 4968 responses total. Due to the unequal sample size and variance between the 54 specifically targeted design ratings and the 4968 average survey responses, Welch's t-test was used to test for statistical significance. The Welch's t-test is a variation from a normal student's t-test that allows for unequal variance and sample sizes between the two populations. Note that the effective degrees of freedom reported here are approximated using the Welch-Satterthwaite equation. For all statistical tests, an alpha level of 0.05 was used ($\alpha = 0.05$).

As described in Sec. 3.3.3 four designs were inserted five times each into the first stage survey, and three times each into the second stage survey. It is hypothesized that participants who are unable to be consistent with ratings in a category across identical vehicle designs presented throughout the study would also not be consistent with the reasoning behind their judgments in that category. Using inconsistent judgments to train the ANN would result in a less accurate model, as well as a less accurate verification survey, so the beneficial effect of removing less consistent participants from the data set was also tested. The average standard deviation of all random survey data gathered for the study across all categories is 1.04, and the average standard deviation of all ratings targeting specific generated designs is 0.87. It was decided that the average between these two standard deviations, a standard deviation of 0.96, would be used as the cutoff threshold for use in filtering out less consistent participants. More specifically, if the average standard deviation between the four cars, presented 8 times each, is above 0.96, the participant would be removed from the corresponding rating category of the study. This filter removed

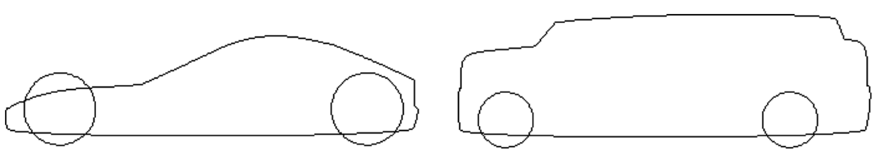

Fig. 3 Generated design most sporty (left) and least sporty (right) for participant 18

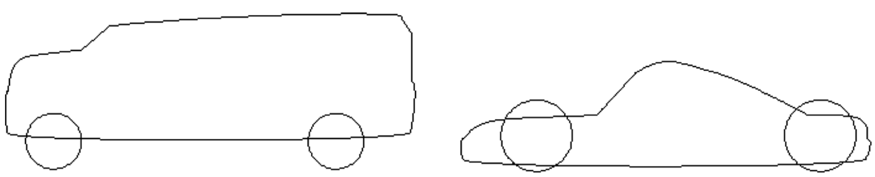

Fig. 4 Generated design most rugged (left) and least rugged (right) for participant 18

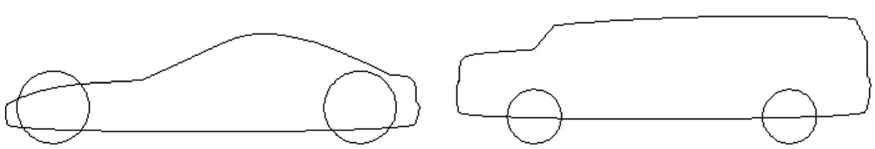

Fig. 5 Generated design most beautiful (left) and least beautiful (right) for participant 18

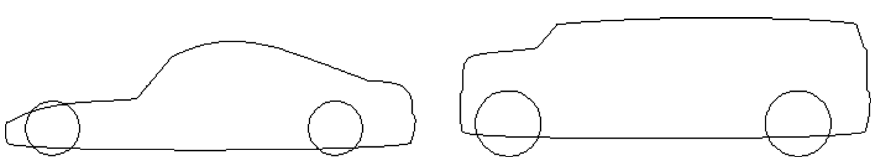

Fig. 6 Generated design most fuel efficient (left) and least fuel efficient (right) for participant 18

one participant from the sportiness category, three participants from the ruggedness category, nine participants from the fuel efficiency category, and six participants from the beauty category. The data for the study are presented both with and without this filter in the results section.

3.5 Survey Results. The average survey ratings from all eighteen participants, without applying the consistency filter, are shown in Fig. 7. In all four categories, the vehicles generated using the ANN and GA learning and generation system to elicit high and low ratings were rated higher and lower than the average of all 4968 ratings in their respective category by statistically significant amounts. The error bars shown in Fig. 7 represent the standard error. Vehicles generated to appear most sporty had significantly higher sportiness ratings than the average ratings in the sportiness category, $t(56.6) = 19.49$, $p < 0.001$, while vehicles generated to appear least sporty had significantly lower sportiness ratings than the average ratings in the sportiness category, $t(61.6) = 24.98$, $p < 0.001$. Similarly, vehicles generated to appear most rugged had significantly higher ruggedness ratings than the average ratings in the ruggedness category, $t(55.0) = 16.38$, $p < 0.001$, while vehicles generated to appear least rugged had significantly lower ruggedness ratings than the average ratings in the ruggedness category, $t(57.2) = 13.58$, $p < 0.001$. Also, vehicles generated to appear most fuel efficient had significantly higher fuel efficiency ratings than the average ratings in the fuel efficiency category, $t(53.7) = 5.38$, $p < 0.001$, while vehicles generated to appear least fuel efficient had significantly lower fuel efficiency ratings than the average ratings in the fuel efficiency category, $t(53.8) = 2.25$, $p = 0.014$. Lastly, vehicles generated to appear most beautiful garnered significantly higher beauty ratings than the average ratings in the beauty category, $t(53.8) = 2.72$, $p = 0.004$, while vehicles generated to appear least beautiful had significantly lower beauty ratings than the average ratings in the sportiness category, $t(54.0) = 6.37$, $p < 0.001$.

The strength of the results differs across aesthetic categories, which is shown by the magnitude of the span between the targeted high and low ratings in each aesthetic category. The strength of the results are strongest for sportiness, followed closely by ruggedness, with somewhat weaker spans observed for ratings of fuel efficiency and beauty. This discrepancy suggests that the methods discussed in this paper are more consistently effective at generating designs that target desired ratings in sportiness and ruggedness than with fuel efficiency and beauty. At this stage it is difficult to determine which step or steps of the process discussed in this paper are responsible for this difference in performance. One possible cause for this difference in performance is that some participants were less consistent in performing judgments of fuel efficiency and beauty than sportiness and ruggedness. The effect of this potential inconsistency is tested next.

A consistency filter was developed to remove participants who were unable to consistently rate the consistency verification

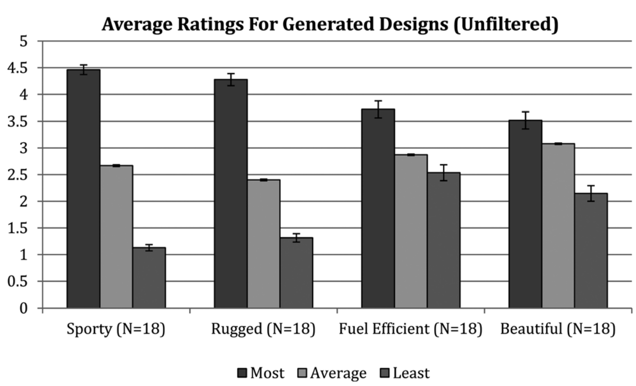

Fig. 7 Average ratings for generated designs, unfiltered

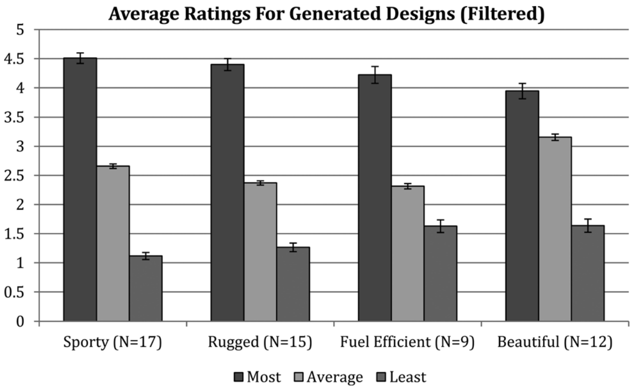

Fig. 8 Average ratings for generated designs, with less consistent participants filtered out

designs in the survey. It is assumed that those participants who were less able to consistently judge the same vehicle design presented multiple times have a more weakly defined opinion about that particular judgment category, and thus it would be difficult to reliably generate vehicle designs that satisfy them in that particular judgment category. All participants who had a standard deviation greater than 0.96 across the consistency verification designs in each category were filtered out with the assumption that a lack of consistency across the test vehicles would also manifest itself as a lack of consistency with the vehicles used to train the ANN, and a lack of consistency in rating the generated vehicles. Applying this filter removed one participant from the sportiness category, three participants from the ruggedness category, nine participants from the fuel efficiency category, and six participants from the beauty category. The average ratings for the generated designs are shown in Fig. 8. When compared to the unfiltered results, the differences between the ratings for the generated designs in comparison with the average ratings in each category were even larger. Note that in addition to changes to the averages of the high and low ratings, the means for the average ratings for each category changed as a result of removing the ratings of the filtered participants from the average.

With the filtered results, the results of the significance tests became even stronger. Vehicles generated to appear most sporty had significantly higher sportiness ratings than the average ratings in the sportiness category, $t(53.7) = 20.14$, $p < 0.001$, while vehicles generated to appear least sporty had significantly lower sportiness ratings than the average ratings in the sportiness category, $t(58.4) = 24.57$, $p < 0.001$. Similarly, vehicles generated to appear most rugged had significantly higher ruggedness ratings than the average ratings in the ruggedness category, $t(46.5) = 19.51$, $p < 0.001$, while vehicles generated to appear least rugged had significantly lower ruggedness ratings than the average ratings in the ruggedness category, $t(48.8) = 14.59$, $p < 0.001$. Also, vehicles generated to appear most fuel efficient had significantly higher fuel efficiency ratings than the average ratings in the fuel efficiency category, $t(27.2) = 13.04$, $p < 0.001$, while vehicles generated to appear least fuel efficient had significantly lower fuel efficiency ratings than the average ratings in the fuel efficiency category, $t(28.1) = 6.19$, $p < 0.001$. Lastly, vehicles generated to appear most beautiful garnered significantly higher beauty ratings than the average ratings in the beauty category, $t(36.2) = 5.92$, $p < 0.001$, while vehicles generated to appear least beautiful had significantly lower beauty ratings than the average ratings in the beauty category, $t(36.7) = 13.20$, $p < 0.001$.

By applying a consistency filter, the strength of the results became consistently strong. Applying this filter removed one participant from the sportiness category, three participants from the ruggedness category, nine participants from the fuel efficiency category, and six participants from the beauty category, showing that as a whole the participants were most consistent at judging sportiness, followed by ruggedness, and then beauty, and lastly fuel efficiency. While this filter has shown that removing less consistent participants from the training process increases the performance of the methods described in this paper, the characteristics and root cause of this inconsistency are not yet understood. Further research is needed to examine improved ways to better survey and model these consumers.

4 Mapping Tradeoffs Between Ruggedness and Aerodynamic Performance

In this section, the tradeoff between perceived ruggedness and aerodynamic performance in highly rugged cars is analyzed with the help of a plot of the Pareto Set. In order to plot the tradeoffs between two objectives, two separate objective functions are needed. To represent judgments of perceived ruggedness, the artificial neural network trained on survey data from Participant 15 from the experiment in Sec. 3 is used as the first objective function. Since the Multi-objective Genetic Algorithm Tool seeks to minimize all objective functions, the negative of the output of the network was used. Note that the absolute value of the ruggedness rating objective function was plotted for clarity. In order to represent aerodynamic performance, a parametric aerodynamic model, described in the following section, was developed for use as the second objective function.

4.1 Aerodynamic Model. In order to assess the accuracy of participant judgments, a method for determining the aerodynamic drag properties of a vehicle's geometry was needed. In current production vehicles, this task is typically performed two ways, wind tunnel testing, and computational fluid dynamics (CFD) analysis. Due to the quantity of models that needed to be assessed, neither of these methods were deemed efficient. Instead, a feature based aerodynamic drag coefficient metamodel called CDaero by Guan, Chan, and Calkins [42–45] was adapted for use with our parametric vehicle model.

Guan's model was based on work by Carr and Stapleford [45], who developed 13 discrete parametric equations that modeled the primary contributions of aerodynamic drag. This model was updated by Guan [43] for use with modern vehicle shapes. Guan's model uses 51 parameters to define vehicle shape and to calculate 13 primary contributions to drag coefficient, and in test cases was reportedly able to predict vehicle drag coefficient to within $+8.2\%$ to -15.2% of actual wind tunnel results. This model was further refined by Chan [44] to yield $\pm6\%$ accuracy on the same test cases.

The vehicle model used in CDaero is more detailed and extensive than our parametric vehicle model. Six of the 13 primary contributions, wheel wells, external mirrors, drip-rails, window recesses, mudflaps, and the cooling system were extraneous to our needs since they modeled vehicle traits that did not exist in our model. Modeling only the seven remaining relevant contributions reduces the list of 51 necessary parameters to only 32 parameters. The vehicle model used in CDaero is three-dimensional, while our model is two-dimensional. Of the 32 parameters, 14 pertain only to width, or details that are irrelevant in our model, and thus were assumed to equal a constant average value, resulting in a simplified CDaero model with only 18 parameters. In order to convert our 12-parameter model into the 18 required parameters, analytical solutions for a number of heights, positions, and points were used. Remaining points that could not be ascertained analytically were found by computationally searching along various vehicle curves for maximum curvature, points of inflection, and maximum and minimum height. The results from these searches at extreme parameter values allowed values to be approximated using interpolation. The vehicle designs in our experiment spanned a wider range of dimensions than was assumed in the CDaero model, thus several equations had to be readapted for model stability. The resulting model allows us to assign an aerodynamic drag coefficient to each vehicle design generated by our parametric model. While this aerodynamic drag coefficient is not absolute, the trends of improved or worsened aerodynamic performance between different vehicle designs are accurately captured. For more detail pertaining to this aerodynamic model and its adaptation, refer to Ref. [39], and for verification of its performance, see Ref. [46].

4.2 Pareto Frontier. The Pareto Set plotted in Fig. 9 was found using the Multi-objective Genetic Algorithm Tool in MATLAB's Optimization Toolbox. All points in this plot represent an optimal design based on a chosen tradeoff between ruggedness and aerodynamic performance. The purpose is to study the tradeoff in cars that are perceived to be highly rugged, so a subset of the data are plotted from ruggedness ratings of 3.8 and higher. It is interesting to note that due to the characteristics of this specific artificial neural network, the plotted Pareto Set appears to be a piecewise response. A closer investigation of the vehicle designs at each of the four marked points suggests that this is because there are two different design families that provide optimal designs in this Pareto plot. In this research, design families are defined as groupings of designs along the Pareto Frontier wherein the parameters vary continuously as you step along the Pareto Frontier and vary your tradeoff between goals. In other words, there are no sudden jumps or discontinuities in the design space, which would indicate a transition to another design family. Starting from the upper right of Fig. 9 with the highest achieved level of ruggedness rating of 6.04 and an aerodynamic coefficient of drag of 0.320 at data point A, the corresponding vehicle design for data point A is part of the first design family, and is shown in Fig. 10(*a*). In traversing from data point A to data point B, there appears to be a smooth transition trading some ruggedness for gains in aerodynamic performance, with diminishing returns of how much aerodynamic performance is gained per unit of ruggedness lost when approaching data point B. The ruggedness for the design at data point B has been reduced to 4.68, while aerodynamic performance has improved to 0.303. The vehicle design at data point B is still part of the first design family, and is shown in Fig. 10(*b*).

In studying the vehicle designs (not shown) between data point A and data point B a smooth and continuous transition between the two designs is seen, with the majority of the change happening as a result of a reduction of body height. As noted previously, these designs belong to a single design family of optimal tradeoffs between ruggedness and aerodynamics that fall along the dotted line traveling through data points A and B in Fig. 9. At data point C a discontinuity in slope is observed. The vehicle design at data point C is part of a second design family, and is shown in Fig. 10(*c*). This transition shows a discontinuous jump from the smooth progression of the vehicle design family observed

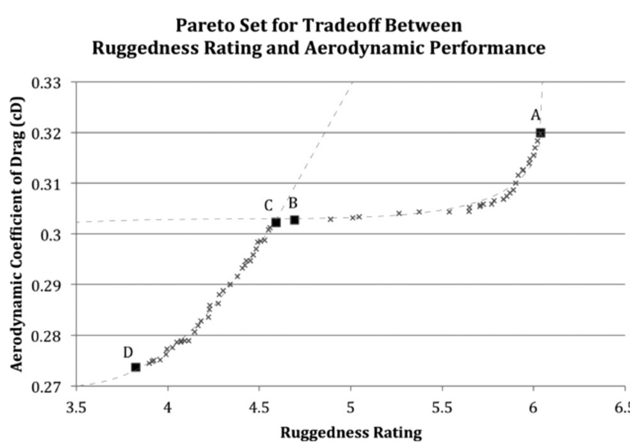

Fig. 9 The Pareto Frontier of optimal tradeoffs between the ruggedness rating and aerodynamic performance

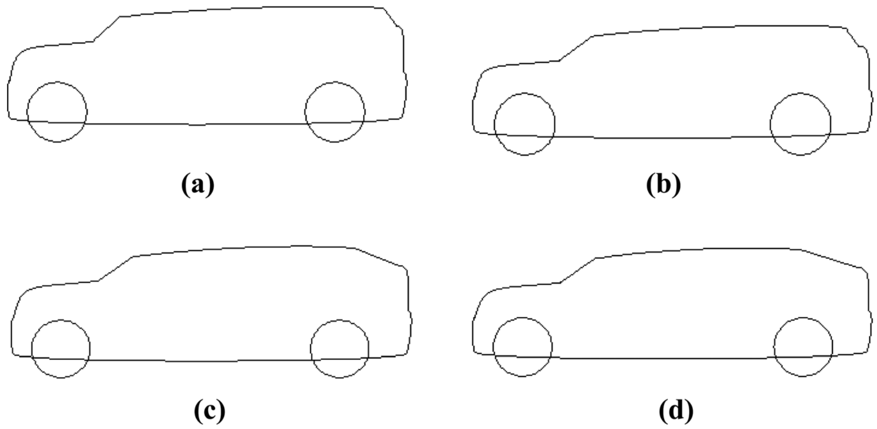

Fig. 10 Pareto optimal solutions for (*a*) Point A, (*b*) Point B, (*c*) Point C, and (*d*) Point D in Fig. 9

between data point A and data point B. The rear windscreen at data point C becomes much more highly raked, and the body height has increased above that of data point B. A closer study of the models indicates that this increased raking of the rear windscreen causes a reduction in perceived ruggedness, but also an increase in aerodynamic performance, resulting in almost exactly the same ruggedness rating and aerodynamic performance as the vehicle design at data point B. Where these two design families (depicted by the two dotted lines) intersect, there are two different but equivalent vehicle designs that have the same ruggedness rating and aerodynamic performance, belonging to each of the two design families. At ruggedness ratings above 4.6, designs in the first design family, which contains data points A and B are more optimal, whereas below ruggedness ratings of 4.6, designs in the second design family, which contains data points C and D are more optimal. The Pareto Set is made up of designs from both design families, but only from regions where the respective design family is optimal. The design at data point C has a ruggedness rating of 4.59 and an aerodynamic coefficient of 0.302. In studying the vehicle designs between data point C and data point D (not shown), which both belong to the second design family, another smooth and continuous transition between the two cars is seen, with the majority of the change once again happening as a result of a reduction of body height, resulting in the vehicle design at data point D, which is shown in Fig. 10(*d*). Once again there is a relationship where perceived ruggedness is being traded for an improvement in aerodynamic performance. In the tradeoff shown between data point C and data point D, the effect of diminishing returns appears not to be as strong as the one seen between data point A and data point B. The perceived ruggedness at data point D is 3.81, and the aerodynamic coefficient is 0.274.

This mapping of the Pareto Set allows designers to better visualize the design options that are afforded to them, and can help designers to make better design decisions, especially pertaining to tradeoffs and synergies between design goals, and identifying different families of designs.

5 Tradeoffs Between Vehicle Volume and Aerodynamic Performance

In the current climate of high fuel prices and people's desires for capacious and useful vehicles, it should come as no surprise that tradeoffs between a vehicle's interior volume and a vehicle's aerodynamic performance are of utmost concern. In Sec. 4, a Pareto Set was generated to map the interface between two tradeoffs. While this method enables a designer to visualize tradeoff possibilities when the desired goals are unknown, the method can be computationally expensive when done on the computer, or resource intensive when done manually. In some design problems a subset of specific target goals are known. This section will focus on designing a series of vehicles with specific performance goals that demonstrate how the ANN and GA methods developed earlier in this paper can be used to target specific design goals. In Sec. 5.1, three vehicles are designed using traditional methods of GA optimization for functional performance. In Sec. 5.2, six vehicles are designed to accomplish the same functional goals as in Sec. 5.1, but with the addition of a stylistic form goal of sportiness, which is assessed using ANNs developed in Sec. 3 of this paper.

In general, GAs require fitness functions that provide a single scalar value for a given set of parameters. This can pose difficulties in multi-objective optimization, where the different objectives can be represented in substantially different ways. There have been several approaches to combine multiple objectives into a single scalar function that is appropriate for use in GAs [47–49]. Weighted sum approaches are one common approach where multiple objectives are adapted into a single fitness function made up of a sum of normalized positive scalar objective functions, which can then be weighted if desired [50,51].

5.1 Optimizing Aerodynamic Performance With Fixed Volume. The designs generated in this section focus on two purely functional goals: actual aerodynamic performance and vehicle volume. More specifically, the goal of this design problem is to generate the most aerodynamic shapes that can be achieved to satisfy a given volume. These two functional goals are assessed using the aerodynamic performance model and vehicle volume models discussed earlier, and do not use the ANN-based stylistic form model. Because the volume is constrained in this case, and the aerodynamic coefficient of drag is to be minimized, the fitness function must be set up accordingly using a weighted sum approach [50,51]. When using a weighted sum approach, the fitness function is comprised of multiple objective functions that are normalized to span a common range of differences between maximum and minimum values, and manual weights for the objective function that can be added to adjust the preferences of the system [52]. The absolute values of the objective functions are unimportant since the difference spanned between the maximum and minimum values of the objective function is normalized. The first objective function in this problem, the aerodynamic coefficient of drag, which normally has a function range of approximately 0.20–0.50 was multiplied by 333 to be normalized to a range of 66–166, a difference of 100. The second objective function is structured differently from the first because the volumetric assessment in this model has a target value. Because of this target value, the objective function should no longer penalize the fitness function after its value drops below the desired target. Since the MATLAB GA function works to minimize objective function, and the goal is to maximize the volume, the objective function used is

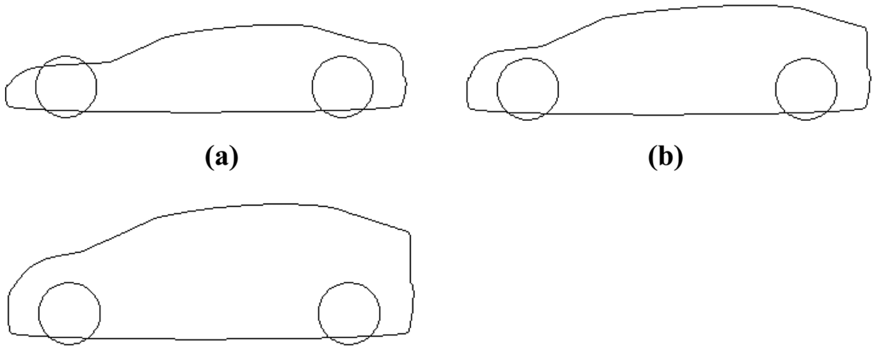

Fig. 11 (*a*) **Most aerodynamic solution, 65 cu. ft., (*b*) most aerodynamic solution, 75 cu. ft., and (*c*) most aerodynamic solution, 85 cu. ft.**

the negative of the volumetric assessment function. The target value is implemented with a Boolean operator, which sets the volumetric assessment function equal to the desired volume when the volume assessment is equal to or larger than the desired volume, thus removing the effect of the objective function once the target is reached. The range of this function is also normalized for a difference between maximum and minimum values of approximately 100. In all generated designs, the constraints were found by inspection to be active.

Using this new fitness function, the GA generated the design pictured in Fig. 11(*a*) as the most aerodynamic shape that can be achieved with a minimum volume of 65 cu. ft. with a coefficient of drag of 0.196. By increasing the targeted minimum volume to 75 cu. ft., the design pictured in Fig. 11(*b*) was designed as the most aerodynamic shape with a coefficient of drag of 0.223. Increasing the targeted minimum volume to 85 cu. ft. results in the design pictured in Fig. 11(*c*) as the most aerodynamic shape with a coefficient of drag of 0.254. Using the methods discussed here, a GA in conjunction with two functional design models is able to generate functional designs that target specific performance targets. One interesting note is that the most aerodynamic car design that can encompass a larger interior volume, as shown in Fig. 11(*c*), appears to resemble the Toyota Prius, a real vehicle that was designed to balance goals of aerodynamic performance and spaciousness.

5.2 Adding Sportiness to a Useful Design. In Sec. 5.1, three vehicles were designed to illustrate the powerful computational design abilities that can be unveiled by coupling a GA with two functional assessment models. The vehicles designed were of a purely functional shape and did not use the ANN-based stylistic form model. The goal of this research is to couple stylistic form and function to allow for synergies and optimal tradeoffs. Leveraging the ANN-based stylistic form model developed in Sec. 3, this section aims to computationally add stylistic sportiness, based on human participant data, to the three designs generated in Sec. 5.1. The ANNs that were trained in Sec. 3 differed greatly between participants. In order to illustrate the range of differences modeled, this section will demonstrate how the GA generates vehicles that add sportiness to the previous set of functional goals (aerodynamic and a target volume) using the ANNs from two different participants for comparison. In Sec. 5.1, an aerodynamic coefficient of drag of 0.254 was achieved for a vehicle with 85 cu. ft. of volume. This represents the optimal aerodynamic performance a vehicle with that volume can achieve under the constraints of our model, without the considerations of stylistic sportiness assumed in this section. In this design problem, the goal is to design the sportiest designs that balance a targeted vehicle volume with an optimal aerodynamic performance that also must perform at least as well aerodynamically ($cD = 0.254$) as the 85 cu. ft. vehicle designed in Sec. 5.1.

In this problem, three different types of objective functions are used. (1) In the first objective function, the sportiness rating, is maximized. For compatibility with MATLAB's GA Optimization Tool, which aims to minimize objective functions, the objective function must be formatted such that the negative of the sportiness rating is minimized. The sportiness rating is assessed through the ANN developed in Sec. 3.2) The second objective function is to maximize the volume of the vehicle until specific target values have been achieved. Much like in Sec. 5.1, the negative of the objective function is to be minimized until the target value is reached, at which point a Boolean operator enforces the target value. 3) The third objective function is the aerodynamic coefficient of drag, which is to be minimized, and must not exceed 0.254. This maximum coefficient of drag is enforced using a penalty function [53], which adds a large penalty value (1000) to the objective function value if the coefficient of drag exceeds 0.254. All objective functions were normalized to a range between maximum and minimum values of 100 using the methods discussed in Sec. 5.1 and were summed together without additional weighting. The ANNs used to model sportiness in this problem are from Participant 3 and Participant 18. The sportiest and least sporty designs for Participant 18 without any other objective functions are shown in Fig. 3. Note that because the ANNs for the two participants are different and have different ranges, the normalization weights used were also different in order to achieve the desired range of 100.

First, two vehicles were designed for a target volume of 65 cu. ft. The design based on the sportiness ANN from Participant 3 is shown in Fig. 12(*a*). This design achieves the target volume of 65 cu. ft., and manages it with a sportiness rating of 7.03, and an aerodynamic performance of 0.211. Similarly, the design based on the sportiness ANN from Participant 18 is shown in Fig. 12(*b*). This design also achieves the target volume of 65 cu. ft., and manages it with a sportiness rating of 5.31 and an aerodynamic performance of 0.202.

Next, two vehicles were designed for a target volume of 75 cu. ft. The vehicle designed using the sportiness ANN from Participant 3 is shown in Fig. 12(*c*). This vehicle achieves the target volume of 75 cu. ft., and manages it with a sportiness rating of 6.86 and an aerodynamic performance of 0.249. The vehicle designed using the sportiness ANN from Participant 18 is shown in Fig. 12(*d*). This vehicle achieves the target volume of 75 cu. ft., and manages it with a sportiness rating of 4.84 and an aerodynamic performance of 0.239.

Lastly, two vehicles were designed for a target volume of 85 cu. ft. The vehicle designed using the sportiness ANN from Participant 3 is shown in Fig. 12(*e*). Due to the restrictions on aerodynamic performance and sportiness, this vehicle is not able to reach its target volume of 85 cu. ft.; instead it measures 79.96 cu. ft. This increase in volume has also reduced the sportiness rating to 5.37 and the aerodynamic performance is at the limit of the penalty function at 0.254. The vehicle designed using the sportiness

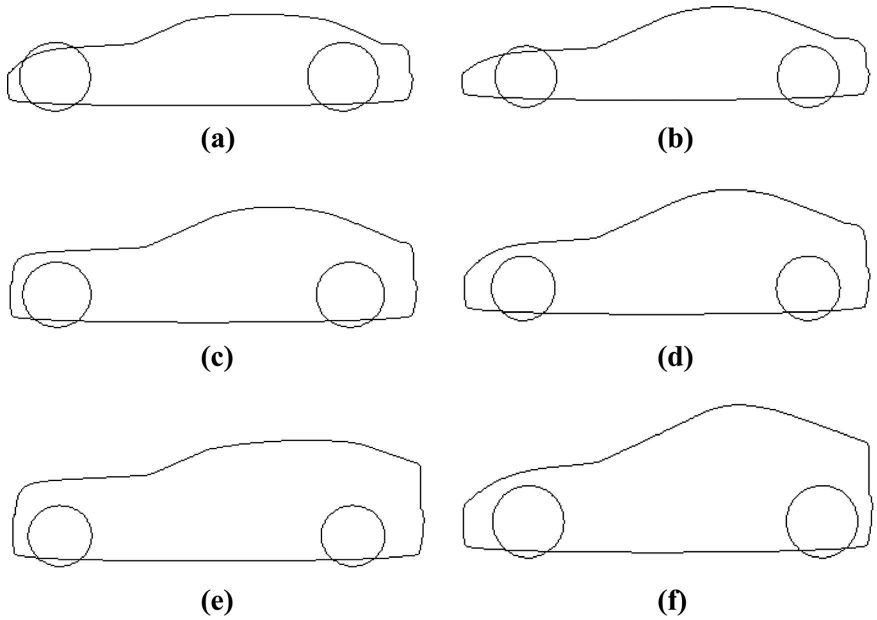

Fig. 12 (*a*) Participant 3 65 cu. ft., (*b*) participant 18 65 cu. ft., (*c*) participant 3 75 cu. ft., (*d*) participant 18 75 cu. ft., (*e*) participant 3 79.96 cu. ft., and (*f*) participant 18 82.59 cu. ft.

ANN from Participant 18 is shown in Fig. 12(*f*). This vehicle also was unable to achieve the volumetric goal of 85 cu. ft., instead achieving only 82.59 cu. ft., while the sportiness rating has dropped to 4.45, and the aerodynamic performance has also been pushed to the limit of the penalty function at 0.254. In both cases, the vehicle designs did not meet the target volume of 85 cu. ft., instead the designs settled on a balance between the needs of the two equally weighted objective functions of volume and sportiness. Had the volume objective also been represented with a penalty function, like the aerodynamic goal, the sportiness objective function would have been dominated by both penalty functions. The resulting vehicle would resemble the vehicle shown in Fig. 11(*c*), a design that did not consider the sportiness objective function, but did achieve both the aerodynamics and volume goals.

In this section six vehicles were designed that balance the functional goals used in Sec. 5.1 with the desire to build sporty vehicles based on learned preferences through the ANN. ANNs from two different participants were used to generate the vehicles in this section. As a result, two strikingly different sets of vehicle designs resulted based on judgments of what each participant feels is the sportiest design. While the introduction of the sportiness objective prevented the designed vehicles from achieving the 85 cu. ft. volumetric target, two uniquely and differently designed vehicles that show optimal tradeoffs with the three goals were generated.

The experiment in this section also helps to confirm that the ANN does capture interaction effects in stylistic form. This is illustrated by studying the position and size of the wheels in each vehicle design. Wheel size and wheel position are not active variables in either the aerodynamic model or the volumetric assessment. As the changes in volume and aerodynamics were enacted on the vehicle designs resulting in shape changes, the sportiest size and positioning of the wheels changed accordingly. If interaction effects were not modeled, the wheels would remain stationary throughout all body shape changes for cars built with each participant's ANN. In checking for the effects of the stochasticity of the GA used, each vehicle was generated multiple times using the GA, and the shift in size and position of the wheels were consistent throughout all trials, showing that the changes were in fact caused by modeled interactions effects in the ANN, and not the stochasticity of the GA.

6 Discussion and Conclusions

This study suggests that it is possible to train an artificial neural network to mimic human judgments of stylistic form. It is also suggested that genetic algorithms can be successfully used to generate designs that rate high and low consistently from both the ANN and human participants. In all eight cases, designs that were generated with the goal of being rated high or low in each of the four categories received ratings that were significantly higher and lower respectively than the average ratings in each category. The system appears to be more accurate at judging and generating vehicles that elicit high and low ratings in sportiness and ruggedness than it is with ratings in fuel efficiency and beauty. It is possible that this is because the defining terms have been more clearly defined by societal norms for sporty and rugged cars than for fuel efficient or beautiful cars.

A consistency verification metric was developed and used to filter the results from this study. This metric removed participants from categories where they were unable to demonstrate sufficient consistency with identical vehicle designs that were presented throughout the surveying process. Applying this filter removed one participant from the sportiness category, three participants from the ruggedness category, nine participants from the fuel efficiency category, and six participants from the beauty category. Removing these participants from these respective categories strengthened the results further by increasing the margin between the targeted high and low ratings and the average nontargeted ratings in each category, especially in the judgment categories of beauty and fuel efficiency. The efficacy of this consistency filter highlights the need for further research to examine ways to better understand and model the desires of consumers who could not be effectively surveyed and modeled by the methods discussed in this paper.

This ANN and GA based framework enables computers to quantify human judgments of stylistic form, which in turn opens the door for computers to treat stylistic form as an objective function in optimization. By plotting the optimal tradeoffs between a stylistic form goal and a functional goal, a Pareto Frontier can be mapped. This mapping of a Pareto Frontier allows designers to better visualize the design options that are afforded to them, and can help designers to make better design decisions, especially pertaining to tradeoffs and synergies between design goals, and identifying different families of designs. This ability to treat stylistic form goals concurrently with functional goals in multi-objective optimization also enables computers to generate designs that reflect not only functional performance, but also specific consumer desired stylistic traits.

In conclusion, it is clear that the framework developed in this paper enables computers to quantify and model human judgments of desired stylistic form, and enables computers to use these models to inform the design process, and generate designs that elicit desired stylistic judgments from consumers, while also incorporating functional performance goals. This method can be applied to a wide range of product categories and can help designers to better manage differing goals of stylistic form and function in the design process. Additionally, this method can be applied to streamline marketing research, it can be extended to act as a metric to objectively rate products based on aesthetic performance, or can be implemented to automatically generate customized product designs that suit the functional and aesthetic desires of specific individuals.

Acknowledgment

This work was supported by the National Science Foundation under Grant Nos. BCS0717957 and SBE0738058.

References

[1] Birkhoff, G. D., 1933, *Aesthetic Measure*, Harvard University Press, Cambridge.
[2] Pye, D., 1969, *The Nature of Design*, Reinhold Books Corporation, New York.
[3] Stiny, G., 1980, "Introduction to Shape and Shape Grammars," Environ. Plann. B, **7**(3), pp. 343–351.
[4] Cagdas, G., 1996, "A Shape Grammar: The Language of Traditional Turkish Houses," Environ. Plann. B, **23**(4), pp. 443–464.
[5] Agarwal, M., and Cagan, J., 1998, "A Blend of Different Tastes: The Language of Coffee Makers," Environ. Plan. B: Plan. Des., **25**(2), pp. 205–226.
[6] Pugliese, M. and Cagan, J., 2002, "Capturing a Rebel: Modeling the Harley-Davidson Brand through a Motorcycle Shape Grammar," Res. Eng. Des., **13**(3), pp. 139–156.
[7] Orsborn, S., and Cagan, J., 2009, "Automatically Generating Form Concepts According to Consumer Preference: A Shape Grammar Implementation With Software Agents," ASME J. Mech. Des., **131**, p. 121007.
[8] Orsborn, S., and Cagan, J., 2009, "Multiagent Shape Grammar Implementation: Automatically Generating Form Concepts According to a Preference Function," J. Mech. Des., **131**, p. 121007.
[9] Hsiao, S. W., and Liu, M. C., 2002, "A Morphing Method for Shape Generation and Image Prediction in Product Design," Des. Stud., **23**(5), pp. 533–556.
[10] Smith, R. C., Pawlicki, R., Kokai, I., Finger, J., and Vetter, T., 2007, "Navigating in a Shape Space of Registered Models," IEEE Trans. Vis. Comput. Graph., **13**(6), pp. 1552–1559.
[11] Hsiao, S. W., and Huang, H. C., 2002, "A Neural Network Based Approach for Product Form Design," Des. Stud., **23**(1), pp. 67–84.
[12] Ranscombe, C., Hicks, B., Mullineux, G., and Singh, B., 2011, "Visually Decomposing Vehicle Images: Exploring the Influence of Different Aesthetic Features on Consumer Perception of Brand," Des. Stud., **33**(4), pp. 319–341.
[13] Hazelrigg, G. A., 1998, "A Framework for Decision-Based Engineering Design," J. Mech. Des., **120**(4), pp. 653–659.
[14] Li, H., and Azarm, S., 2000, "Product Design Selection Under Uncertainty and With Competitive Advantage," J. Mech. Des., **122**(4), pp. 411–418.
[15] Wassenaar, H. J., Chen, W., Cheng, J., and Sudjianto, A., 2005, "Enhancing Discrete Choice Demand Modeling for Decision-Based Design," ASME J. Mech. Des., **127**(4), pp. 514–523.
[16] Michalek, J. J., Ceryan, O., Papalambros, P. Y., and Koren, Y., 2006, "Balancing Marketing and Manufacturing Objectives in Product Line Design," ASME J. Mech. Des., **128**(6), pp. 1196–1204.
[17] Sylcott, B., Cagan, J., and Tabibnia, G., 2011 Pending, "Understanding of Emotions and Reasoning During Consumer Tradeoff Between Function and Aesthetics in Product Design," Proceedings of the ASME 2011 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference.
[18] Norman, D., 2004, *Emotional Design: Why we Love (or Hate) Everyday Things*, Basic Books, New York.
[19] Govers, P., Hekkert, P., and Schoormans, J. P. L., 2003, *Happy, Cute and Tough: Can Designers Create a Product Personality that Consumers Understand*?, CRC Press, Boca Raton, FL, pp. 345–349.
[20] Nagamachi, M., 2002, "Kansei Engineering as a Powerful Consumer-Oriented Technology for Product Development," Appl. Ergon, **33**, pp. 289–294.
[21] Bouchard, C., Lim, D., and Aoussat, A., 2003, "Development of a Kansei Engineering System for Industrial Design: Identification of Input Data for KES," 6th Asian Design International Conference, ADC, Tsukuba.
[22] Osgood, C., Suci, G., and Tannenbaum, P., 1957, *The Measurement of Meaning*, University of Illinois Press, Urbana, Il.
[23] Reid, T., Gonzalez, R., and Papalambros, P., 2009, "A Methodology for Quantifying the Perceived Environmental Friendliness of Vehicle Silhouettes in Engineering Design," Proceedings of the ASME 2009 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference.
[24] Achiche, S., and Ahmed, S., 2008, "Mapping Shape Geometry and Emotions Using Fuzzy Logic," Proceedings from ASME 2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference.
[25] Yannou, B., Dihlmann, M., and Awedikian, R., 2008, "Evolutive Design of Car Silhouettes," Proceedings from ASME 2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference.
[26] Kim, H. S., and Cho, S. B., 2000, "Application of Interactive Genetic Algorithm to Fashion Design," Eng. Applic. Artif. Intell., **13**(6), pp. 635–644.
[27] Yanagisawa, H., and Fukuda, S., 2004, "Development of Interactive Industrial Design Support System Considering Customer's Evaluation," JSME Int. J., **47**(2), pp. 762–769.
[28] Takagi, H., 2001, "Interactive Evolutionary Computation: Fusion of the Capabilities of EC Optimization and Human Evaluation," Proc. IEEE, **89**(9), pp. 1275–1296.
[29] Mitchell, T., 1997, *Machine Learning*, McGraw Hill, NewYork.
[30] Fukushima, K., 1988, "Neocognitron: A Hierarchical Neural Network Capable of Visual Pattern Recognition," Neural Networks, **1**(2), pp. 119–130.
[31] Fukushima, K., 2003, "Neocognitron for Handwritten Digit Recognition," Neurocomputing, **51**, pp. 161–180.
[32] Yasuda, K., Furuta, H., and Kobayashi, T., 1995, "Aesthetic Design System of Structures Using Neural Network and Image Database," Proceedings of the 3rd International Symposium on Uncertainty Modelling and Analysis.
[33] Baluja, S., Pomerleau, D., and Jochem, T., 1994, "Towards Automated Artificial Evolution for Computer-Generated Images," Connection Sci., **6**(2&3), pp. 325–354.
[34] Griffith, N., and Todd, P. M., eds., 1999, *Musical Networks: Parallel Distributed Perception and Performance*, MIT Press/Bradford Books, Cambridge, MA.
[35] Biles, J. A., Anderson, P. G., and Loggi, L. W., 1996," Neural Network Fitness Functions for a Musical IGA," The International ICSC Symposium on Intelligent Industrial Automation and Soft Computing.
[36] Bull, L., 1999, "On Model-Based Evolutionary Computation," Soft Comput., **3**, pp. 76–82.
[37] Tsutsumi, K., and Sasaki, K., 2008, "Study on Shape Creation of Building's Roof by Evaluating Aesthetic Sensibility," J. Math. Comput. Simul., **77**(5–6), pp. 487–498.
[38] Stiny, G., and Gips, J., 1978, *Algorithmic Aesthetics*, University of California Press, Berkley.
[39] Tseng, I. H., 2011, "The Unification of Stylistic Form and Function," Ph.D. thesis, Carnegie Mellon University, PA.
[40] Levenberg, K., 1944, "A Method for the Solution of Certain Non-Linear Problems in Least Squares," Q. Appl. Math., **2**, pp. 164–168.
[41] Marquardt, D. W., 1963, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," J. Soc. Ind. Appl. Math., **11**(2), pp. 431–441.
[42] Calkins, D. E., and Chan, W. T., 1998, "CDaero – A Parametric Aerodynamic Drag Prediction Tool," International Congress and Exposition, SAE Technical Paper 980398, Feb. Detroit, MI.
[43] Guan, L., 1995, "Feature Based Aerodynamic Drag Coefficient Metamodel," M.S. thesis, University of Washington, Seattle, WA.
[44] Chan, W. T., 1997, "Empirical Prediction of Automobile Drag Coefficient," M.S. thesis, University of Washington, Seattle, WA.
[45] Carr, G. W., and Stapleford, W. R., 1981, "A Proposed Empirical Method for Predicting the Aerodynamic Drag of Cars," Motor Industry Research Association (MIRA), England.
[46] Tseng, I. H., Cagan, J., and Kotovsky, K., 2011, "Form Function Fidelity," International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Washington, DC.
[47] Fonseca, C. M., and Fleming, P. J., 1995, "An Overview of Evolutionary Algorithms in Multiobjective Optimization," Evol. Comput., **3**(1), pp. 1–16.
[48] Zitzler, E., and Thiele, L., 1999, "Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach," IEEE Trans. Evol. Comput., **3**(4), pp. 257–271.
[49] Coello, C. A., 2000, "An Updated Survey of GA-Based Multiobjective Optimization Techniques," ACM Comput. Surv., **32**(2), pp. 109–143.
[50] Zadeh, L., 1963, "Optimality and Non-Scalar-Valued Performance Criteria," IEEE Trans. Autom. Control, **8**, pp. 59–60.
[51] Goicoechea, A., Hansen, D., and Duckstein, L., 1982, *Multiobjective Decision Analysis With Engineering and Business Applications*, John Wiley and Sons, NJ.
[52] Marler, R. T., and Arora, J. S., 2004, "Survey of Multi-Objective Optimization Methods for Engineering," Struct. Multidiscip. Optim., **26**, pp. 369–395.
[53] Marler, R. T., and Arora, J. S., 2009, "The Weighted Sum Method for Multi-Objective Optimization: New Insights," Struct. Multidiscip. Optim., **41**(6), pp. 853–862.