Res Eng Design (2008) 18:163–180
DOI 10.1007/s00163-007-0036-8

ORIGINAL PAPER

Identifying product shape relationships using principal component analysis

Seth Orsborn · Peter Boatwright · Jonathan Cagan

Received: 1 September 2006 / Accepted: 29 May 2007 / Published online: 10 October 2007

Abstract Product forms with multiple features, like automobiles, have traditionally accepted feature definitions and relationships between those features. These relationships drive how the product is created by focusing on expected, and accepted, feature development to push the form outside the traditional bounds. This paper uses principal component analysis to determine the fundamental characteristics within vehicle classes. The results of this analysis can then be considered by product designers to create new designs based upon the derived shape relationships. These new designs will be novel due to the non-traditional grouping of characteristics.

Keywords Vehicle design · Product design · Principal component analysis · Shape relationships

1 Introduction

Vehicle designers, like all trained product designers, use their skill and intuition to bring potential designs to reality. The designers manipulate current forms and create new forms to differentiate their product from existing products. The understanding of vehicle differentiations is often based upon a general understanding of the vehicle form and established relationships between vehicle characteristics. Vehicle characteristics that are physically close to each other often have an established relationship, e.g. the hood and front windshield are related through the cowl.

Efforts to scientifically examine what some consider to be artistic expression, the design of the form of the product, has begun to emerge. Biederman (1987) discusses how humans perceive shapes and decompose them into recognizable chunks. McCormack and Cagan (2002) simplified the representation of complex shapes by decomposing the shapes into mathematically simpler forms. Complex curves are decomposed into distinct subshapes composed of straight lines. Groupings of curves are decomposed into simple shapes of triangles. Prats et al. (2006) take this methodology a step further and use it to understand simplified shapes with regard to style and product design. In this paper we look not at the intuitive, or learned, decomposition of a complex shape, but at the statistical decomposition of a complex shape (a vehicle) into fundamental chunks. We do not consider how to take a complex shape and represent it with a simple shape, but how to break a complex shape into chunks of simpler shapes.

Specifically, we use principal component analysis to identify element groupings that differentiate models within a product class. Principal component analysis is a statistical method whereby an original set of related data is reduced in dimensionality to match the original data as closely as possible (Patel et al. 2006). For example, bicycles can be described by many different dimensions: wheel base, head tube angle, bottom bracket height, etc. Each of these dimensions is described using a numerical variable. Principal component analysis finds a set of new variables that are weighted averages of the original variables, where the new set of variables is of smaller dimension than the original set. For instance, the original set of bicycle variables

S. Orsborn · J. Cagan (✉)
Department of Mechanical Engineering,
Carnegie Mellon University, 5000 Forbes Avenue,
Pittsburgh, PA 15213, USA
e-mail: cagan@cmu.edu

P. Boatwright
Tepper School of Business, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, PA 15213, USA

may be dimensions of wheel base, head tube angle, bottom bracket height, and many others, while the new set may be bicycle height and length. The principal components are ordered by the extent of variation that they capture. The first principal component is a vector that captures more of the variation between the original variables than any other principal component, the second component captures the second most variation, and so forth. Each variable is then given a weight related to the vector. The weight describes the influence of that variable on the principal component. For example, the variable that fundamentally describes a bicycle's size may be the seat tube length, the measurement often referenced when describing a bicycle. The seat tube length would have the highest weight for the first principal component. Based upon the weights of all the variables a percentage variance explained is calculated for the first principal component. The second principal component is described using a vector orthogonal to the first principal component. The analysis is repeated, returning different weights for the variables and a smaller percentage variance explained. In many simpler applications only a few principal components account for much of the differences in the objects and identify key elements that distinguish the objects. In this paper principal component analysis is used to analyze similarities and differences in vehicles.

Orsborn et al. (2006) analyzed the exterior form of a sample of three types of existing vehicle classes. It was determined, in conjunction with a vehicle designer, which vehicle characteristics were most important for capturing the general form of the vehicle so that a potential consumer would recognize different vehicles and designers would have the freedom to develop new forms. The forms for fifteen coupes, twenty SUVs, and seven pickups were captured using four-control-point Bezier curves. The data collected from Orsborn et al. (2006) is used as the foundation for the work described in this paper. Rather than using traditional designer heuristics to determine vehicle characteristics, this paper uses principal component analysis to determine the fundamental characteristics within the individual vehicle classes of coupe, pickup, and SUV. These fundamental characteristics at times differ from the traditional and expected feature definition, resulting in a deeper understanding of what features of an existing set of designs must differentiate one form from another. The results from this analysis can then be used by designers to create new vehicles. The designers can take the statistically based curve chunking found in this paper as a foundation for new designs, rather than rely upon a traditional chunking of curves according to vehicle characteristic and spatial relationships, with the understanding that the relationship between these curve chunks will have a strong effect on the gestalt of the vehicle design.

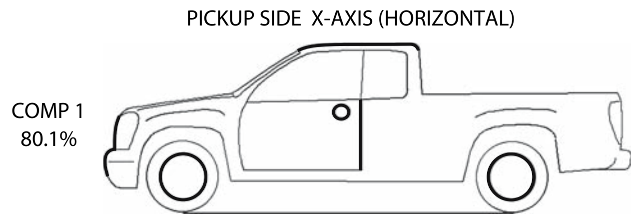

Fig. 1 Chunking of vehicle shapes

One of the results of the analysis was that the complex set of curves that comprise the vehicle form can be efficiently summarized by the principal component vectors. For example, in Fig. 1 it is seen that 80% of pickup differences in this view can be described through a combination of the roof, the front of the nose, the door handle, the back of the door, and the rims. These are not the vehicle characteristics in their entirety, but a chunking of specific curves that together account for 80% of the uniqueness in a pickup. The combination of the results from the statistical measure with an experienced product designer unleashes the full potential of the information through the statistics and the designer working in chorus. Although the focus of this paper is on vehicle design, the methods developed here are applicable to any class of physical products based on a consistent form language. A consistent form language, in the context of this work, is when each product in a product class shares similar shapes or features and the relationship between said shapes is consistent (Biederman 1987). Each product is essentially a parametric manipulation of another product, though some shape may be parametrically manipulated to zero to "remove" the shape.

2 Vehicle data description

In Orsborn et al. (2006), vehicles were chosen from three known classes: coupes, pickups, and SUVs. The selection requirements were that each vehicle have an available blueprint that included the front, side and rear views. Each of the views must be isometric (or as parametrically close as possible) and the three views should complement each other parametrically, i.e. the proportions in each view of the drawing is consistent with the actual vehicle. Table 1 lists the sample vehicles for each class.

To maintain a reasonable degree of homogeneity within classes, all the vehicles chosen were from the 2003 model year. All the pickups chosen were extended cab models. The coupes were all standard coupes. Any vehicles considered to be anomalies in their class were not chosen, such as the Volkswagen Beetle. Anomalies were considered such if they had features that were unique with respect to, or parametrically inconsistent with, the rest of the class.

Table 1 Vehicle sample

	Coupes	SUVs	Pckups
1	Acura RSX	Acura MDX	Chevrolet Silverado
2	Audi TT	BMW X3	Dodge Ram
3	BMW M3	BMW X5	Ford F150
4	Chevrolet Cavalier	Chevrolet Suburban	Ford Ranger
5	Dodge Stratus	Ford Escape	GMC Canyon
6	Ferrari 456M	Ford Excursion	GMC Sonoma
7	Ferrari 612 Scaglietti	Ford Expedition	Toyota Tacoma
8	Ford Mustang	Ford Explorer	
9	Honda Accord	Hyundai Santa Fe	
10	Honda Civic	Kia Sportage	
11	Hyundai Tiburon	Land Rover Free Lander	
12	Mercedes Benz C	Land Rover Range Rover	
13	Mercedes Benz CLK	Mazda Tribute	
14	Mitsubishi Eclipse	Mercedes Benz ML	
15	Toyota Celica	Mitsubishi Montero	
16		Mitsubishi Montero Sport	
17		Porsche Cayenne	
18		Suzuki Grand Vitara	
19		Toyota Land Cruiser	
20		Toyota RAV4	

The Infiniti FX35, while technically an SUV, has an unusual form with respect to most SUVs (Csere 2003), and was not included in the sample.

For the model year 2003, 31 coupe models were sold in the US market. The sample (Orsborn et al. 2006) is composed of 15 of these vehicles, accounting for 48% of the coupes. The pickup sample contains 7 vehicles, which accounts for 50% of the 14 pickup models sold. Fifty-three traditional SUV models were produced in 2003. Of these, ten were identical in form to other SUVs due to rebadging, e.g. the GMC Yukon and the Chevrolet Tahoe. The true population size, with respect to form, was 43 SUVs. Of these, 20 were included in the sample, accounting for 47% of the population. The sample set is summarized in Table 1. Several of the vehicles outside of the sample were used as a control to verify the effectiveness of the analysis.

The intention of this paper is to understand the foundational forms in vehicle classes. In order to easily analyze the form of the vehicle, it must be simplified. In this instance the vehicle is represented by sets of curves. Each set of curves describes a certain characteristic of the vehicle. A minimum number of curves is used to effectively represent the characteristic. Each curve is represented by a four control point Bezier curve. The only exception is the tires and rims, which are represented with circles. It was decided, for the sake of comparison across vehicle classes, that each vehicle should contain every characteristic.

The vehicle characteristics from Orsborn et al. (2006) were used, and will be discussed in detail in this section. Through discussion with a vehicle designer, it was determined that the following vehicle characteristics are the most relevant for sufficiently describing the form of the vehicle (Refer to Fig. 2 for their locations on the vehicle.). The front (1) and rear (2) wheels are each composed of two concentric circles: one for the rim and one for the outside of the tire. The wheel wells (3 and 4) are the sets of curves surrounding the wheels indicating the gap between the

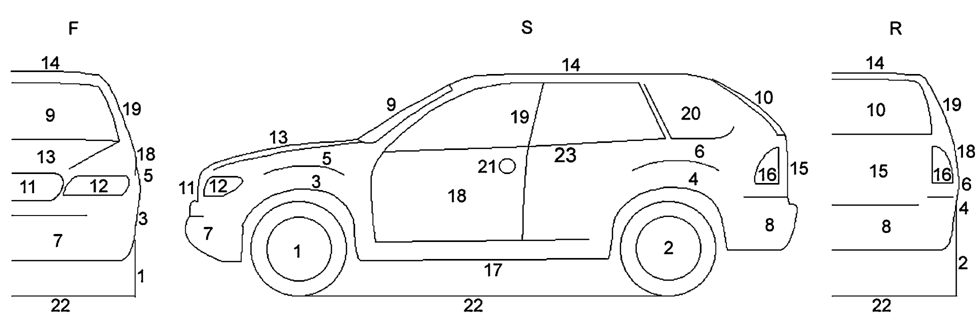

Fig. 2 Vehicle characteristics

wheels and the body of the vehicle. The fender (5 and 6) is the set of curves directly above the wheel well and represents a crease, or sharp change in curvature, commonly on the side of the vehicle. The front bumper (7) includes the curves that connect the grill (11) to the front wheel well. This includes horizontal lines that represent common creases. In some vehicles (like coupes), the grill can be located within the front bumper. In this case, the front bumper connects directly to the hood. Likewise, the rear bumper (8) connects the trunk to the rear wheel well. The front and rear windshields (9 and 10) represent the glass in the front and rear of the vehicle. The headlight (12) generally sits next to the grill (11) between the hood (13) and the front bumper. The roof (14) connects all the glass at the top of the vehicle. The taillight (16) sits within the trunk (15), which is between the rear windshield and the rear bumper. The rocker (17) is the horizontal line between the front and rear wheel wells. The door (18) is represented by key shut lines in the side view and a single curve in front and rear views that connect the fender to the side window. The front side window (19) is located directly above the door with the rear side window (20), typically only found in SUVs and station wagons, located aft of it. The door handle (21) is within the door. The ground (22) is indicated for positioning the wheels and to give a sense placement. The belt line (23), which starts at the bottom of the A-pillar and runs along the bottom of the side windows to the trunk, is an important characteristic. There is no specific curve for the belt line, but it will be built using a combination of the related characteristics: the hood, side windows, and trunk.

Each vehicle class—coupes, pickups, and SUVs—was analyzed individually due to the fact that some vehicle classes have shapes that are indicative of that class. The differences between classes was not the objective, but the foundational form for each vehicle class was the goal of the analysis. The classes were separated into their views (front, rear, side) and dimensions (horizontal and vertical): *y*- and *z*-coordinate forfront and Rear views (horizontal and vertical, respectively), and *x*- and *z*-coordinate for Side view (horizontal and vertical, respectively). The views and the dimensions were separated so that unrelated dimensions would not be analyzed together. As will be seen, many of the same curves are foundational regardless of the view or the dimension. The vehicles were assumed to be symmetric about the longitudinal vertical plane. Due to this, many horizontal curves in the front and rear views have a *y* value of zero. All the curves were normalized according to their matching curves across the vehicle class, to prevent the analysis from being skewed towards larger shapes like hoods and roofs (Patel et al. 2006). For example, in the front view, *y*-coordinate for coupes, all the inner headlight curves were normalized against each other.

3 Chunking key characteristics through principal component analysis

It was shown in Orsborn et al. (2006) that the distinction between vehicle classes can be determined by the parametric ranges of the curves that compose characteristics shared across all vehicle classes. In this section we use principal component analysis to determine which characteristics compose the foundation for a class of vehicles. Through this principal component analysis we are able to chunk together the characteristics that form the foundation of a vehicle form. These chunks can then be used to intentionally create novel vehicles.

The motivation for using principal components analysis is that it provides a method to project a larger dimensional set of data into a smaller set of dimensions, where the smaller set of dimensions captures as much information content as possible from the original space. In mathematical terms, the original data can be represented by $n \times p$ matrix $\mathbf{X}$, where n is the number of observations in the data and p is the number of variables (dimensions). The goal is to define a smaller set of variables or dimensions that approximates the data, an r dimensional subspace where $r < p$. With principal components analysis, the new set of variables $Y_1, Y_2,\ldots,Y_r$ are linear combinations of the original data set X, where $\mathbf{Y}_i = \mathbf{a}_i^T\mathbf{X}$. To preserve as much information content about the original data as possible, the $\mathbf{a}_i$ are selected such that they maximize the sample variance in the new space $\mathbf{a}^T\mathbf{S}\mathbf{a}$, where $\mathbf{S}$ is the sample covariance matrix of X, $\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{X}_i - \mathbf{X})(\mathbf{X}_i - \mathbf{X})^T$ and $\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}$. The first principal component is defined as the linear combination $\mathbf{Y}_1 = \mathbf{a}_1^T\mathbf{X}$ of the original variables that maximizes the value $\mathbf{a}_1^T\mathbf{S}\mathbf{a}_1$, subject to the constraint $\mathbf{a}_1^T\mathbf{a}_1 = 1$ for reasons of identification. The solution to this problem is simply the eigenvector $\mathbf{a}_1^T$ that corresponds to the largest eigenvalue of $\mathbf{S}$ (Krzanowski 1988). The second principal component is similarly defined and must maximize the value $\mathbf{a}_2^T\mathbf{s}\mathbf{a}_2$ subject to the constraints $\mathbf{a}_2^T\mathbf{a}_2 = 1$ and $\mathbf{a}_2^T\mathbf{a}_1 = 0$, in that $\mathbf{a}_2$ is orthogonal to $\mathbf{a}_1$. Similar to the solution for $\mathbf{a}_1$, the solution is the eigenvector $\mathbf{a}_2$ that corresponds to the second largest eigenvalue of $\mathbf{S}$. The pattern continues: the *j*th principal component is the linear combination $\mathbf{Y}_j = \mathbf{a}_j^T\mathbf{X}$, where $\mathbf{a}_j$ maximizes the sample variance subject to constraints $\mathbf{a}_j^T\mathbf{a}_j = 1$ and $\mathbf{a}_j^T\mathbf{a}_j = 0$ $(i < j)$. In such a fashion, the variance of the *j*th component is maximized after accounting for all previous components.

A description of the principal component analysis and the method of extraction of foundational shapes will now be discussed with an active example: the analysis of the front view, *y*-coordinate (horizontal) for coupes, which is seen in the top of Fig. 3. The principal components were found for each coordinate, view, and vehicle class. The first principal component simplifies the description of the data

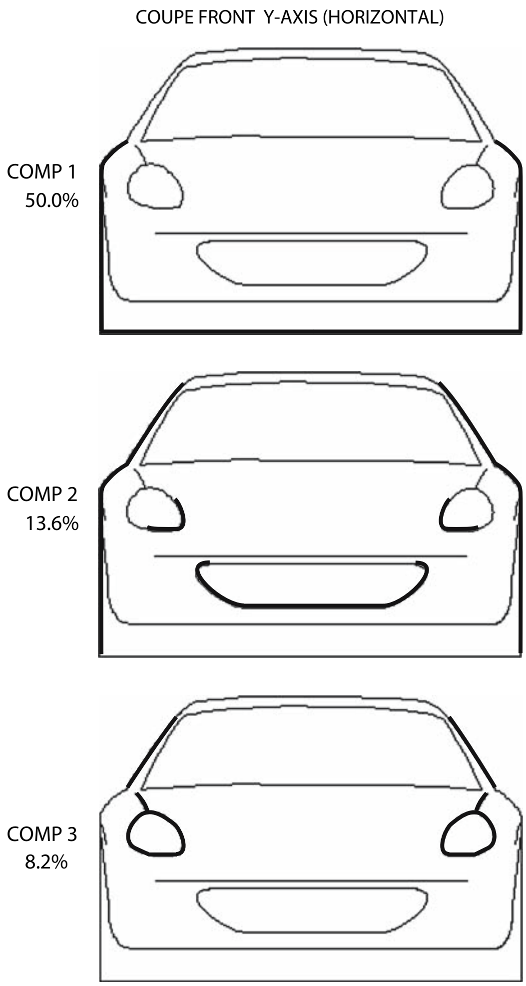

Fig. 3 Coupe front *Y*-axis principal components

through a vector. The data analyzed here is the normalized value of a control point for a specific curve that is part of a characteristic $\overline{\mathrm{cp}}_i = \frac{\mathrm{cp}_i - \mu_{\mathrm{cp}}}{\sigma_{\mathrm{cp}}}$. Each control point has a parametric value, cp_i. The mean of all the related values across the vehicle class, μ_{cp}, is subtracted from the parametric value, which is then divided by the standard deviation, σ_{cp}. For example, the normalized value for control point one, $\overline{\mathrm{cp}_1}$, of the headlight curve 1 for the BMW X_3 is found by taking the original parametric value, subtracting the mean of all the parametric values for control point one of headlight curve 1 and then dividing by the standard deviation of the same. Through the principal component analysis, each control point is assigned a weight based upon its relationship to the principal component vector, the lower the absolute value of the weight the closer to the descriptive vector. For example, door curve 2 control point 1 was assigned a weight of 0.139, the highest value of any control point. This means that the first principal component is being described firstly by this control point. The second highest weight, 0.136, is assigned to the first control point for the third curve of the front fender. This continues on for every single one of the 388 control points. But, a single control point is not enough to determine whether a curve should be included due to the fact that many of the weights are within a percentage of each other. Because a curve is defined by four control points, it was determined beforehand that for a curve to be included in the chunking, at least two-control points from that curve must be included. It was found through the analysis that an isolated control point rarely appeared; for most important curves, the weights on the individual control points were of similar magnitudes.

Since only the dominant curves are of interest, only the top 10% of the weights with the highest absolute value are considered. The control points associated with these weights are then considered, like the two mentioned above. As mentioned previously, it was found that in the top 10%, most of the control points for a specific curve would be found. To continue the example, the first principal component for the front view *y*-coordinate of coupes was dominated by the following curves: Door 1 (upper part of front of door), Door 2 (lower part of front of door), Door 3 (bottom of door),Front Wheel Well 2 (front part of wheel well), Front Fender 1 (front fender crease), Front Fender 3 (outside of fender), Ground 2 (track width), Front Tire and Front Rim. From the front view (top of Fig. 3) these curves all blend together. They can then be summarized by a single aspect of the form: the width of the body and the track. In some principal components, as will be discussed later in detail, the curves do not group together so intuitively. These provide insight into non-obvious shape relationships, which are used as the foundation for describing the vehicle classes.

Each principal component describes a percentage of the variation associated with the original variables. In our example, the first principal component explains 50% of the variation between the coupes in the front view and horizontal coordinate. This has been summarized as the width of the body, which means that the width of the body and track of the vehicle accounts for 50% of the variation between coupes in the front view horizontal coordinate.

This evaluation was done for each coordinate, view, and vehicle class, a total of 18 separate analyses. The number of principal components found for each was as follows: 14 for each of the coupe views and coordinates, 6 for pickups, and 19 for SUVs. The percentage of variance explained was calculated for each principal component. It was determined which principal components dominated based upon Kaiser's Criterion (Kaiser and Rice 1974). According to Kaiser's Criterion, the threshold for relevant principal components is determined by the percentage of variance explained. If the total number of principal components is *X*, then principal components that are retained must describe at least $(1/X) \times 100\%$ of the variation between the original variables. For example, with coupes, front view, y-coordinate there were fourteen principal components. But, after three principal components the percent variance explained dropped below $1/14 \times 100\% = 7.1\%$, the threshold determined by Kaiser's Criterion.

Vehicles are complicated shapes composed of many curves that can be formed in innumerable ways. The principal components capture this variation. The high percentage of variance explained indicates the success of this method in portraying the multitude of curves with a simplified principal component. A summary of the dominant principal components, their percent influence, and composition is itemized in Tables 2, 3, 4, 5, 6, 7, 8, 9 and 10, and shown visually in Figs. 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 and 18. These will be discussed in detail in the following sections.

4 Results of principal component analysis

4.1 Results for coupes

As indicated in Table 2, the first three principal components for the front view y-coordinate (horizontal) of coupes explain 71.8% of the variance and are composed of curves shown in Fig. 3. In Fig. 3, 50.0% of the differences

Table 2 Principal components of coupes front

	71.8%	Coupe front Y	79.2%	Coupe front Z
COMP 1	50.0%	Width of body, track	41.0%	Top and bottom of grill, top of roof, cowl, horizontal front bumper line, top of front windshield
COMP 2	13.6%	Width of body and greenhouse, outside and bottom of grill, inside and bottom of headlight	11.8%	Shoulder, wheel size, inside and top and bottom of head light, horizontal front bumper line, outside of fender
COMP 3	8.2%	Headlight, inner hood line, width of greenhouse	10.6%	Horizontal front bumper line, bottom and top and outside of head light, top of roof, top of fender, top of front windshield, bottom of grill
COMP 4			8.4%	Bottom and top of head light, bottom of front bumper, width of body, cowl, wheel size
COMP 5			7.4%	width of greenhouse, head light, width of body, wheel size, cowl, horizontal front bumper line

Table 3 Principal components of coupes side

	77.1%	Coupe side X	67.5%	Coupe side Z
COMP 1	59.6%	Bottom and back of trunk, bottom and top of rear bumper, back and front of rear wheel well, bottom and inside of tail light, top of front windshield	33.9%	Top of roof, top of windows, top of grill, bottom of windshields, beltline
COMP 2	17.5%	Grill, front and horizontal line and top of front bumper, bottom and outside of head light, front of hood, front of front wheel well	15.5%	Rear bumper horizontal line, top and bottom and outside of tail light, horizontal front bumper line, bottom and top and outside of trunk, bottom and top of rear bumper
COMP 3			9.7%	Horizontal line and bottom and back of trunk, top and inside and bottom of headlight, beltline, rocker, horizontal line of front bumper, front of door
COMP 4			8.4%	Horizontal line and bottom of rear bumper, rear of door, beltline, bottom of head light, handle, cowl, bottom of tail light, bottom of front bumper

Table 4 Principal components of coupes rear

	71.6%	Coupe rear Y	74.5%	Coupe rear Z
COMP 1	50.9%	Width of body, fender, track, wheel size	40.2%	Top of roof, top of side window, bottom and top of rear windshield, beltline, top and bottom of rear bumper, top of tail light
COMP 2	13.3%	Width of greenhouse, rear bumper horizontal line, outside of rear windshield, width of body, tail light	12.9%	Horizontal line and top of rear bumper, shoulder, rear fender
COMP 3	7.4%	Top and outside and inside of tail light, width of greenhouse, outside of roof, outside of rear windshield, rear bumper horizontal line, width of body	11.3%	Trunk horizontal line, rear bumper horizontal line, top of roof, wheel size, outside and bottom of tail light, shoulder
COMP 4			10.1%	Horizontal line and bottom of rear bumper, width of greenhouse, width of body, bottom of rear windshield, bottom of tail light

Table 5 Principal components of pickups front

	96.3%	Pickup front Y	76.2%	Pickup front Z
COMP 1	71.9%	Width of body, inside and outside of head light, outer hood, fender	76.2%	Width of greenhouse, cowl, top of roof, top of front windshield, top of grill
COMP 2	24.4%	Width of body, width of greenhouse, cowl, outside of front windshield, track, wheel size, outer hood line		

Table 6 Principal components of pickups side

	80.1%	Pickup side X	80.7%	Pickup side Z
COMP 1	80.1%	Back of door, grill, top and front of front bumper, handle, rim size, top of roof	80.7%	Rear fender, front fender, cowl, top of front windshield, top of side window, handle, top of roof, bed and gate

Table 7 Principal components of pickups rear

	93.8%	Pickup rear Y	72.0%	Pickup rear Z
COMP 1	93.8%	Width of greenhouse, width of body, fender, outside and top of tail light, top of trunk	72.0%	Width of greenhouse, top of roof, top and back of trunk, bottom of rear windshield

between the coupes in the front view and *y*-coordinate was accounted for in principal component 1 (Comp 1), composed of the body width and track width (highlighted with thicker lines). An additional 13.6% is accounted for in Comp 2, composed of the width of the greenhouse and body, the inside and bottom of the headlight and all except the top of the grill. While the width of the body and the greenhouse may be intuitive, the relationship between only part of the headlight and the grill is not obvious. Figure 4 shows the Comp 2 curve chunking on four of the sample vehicles: Audi TT, Ferrari 456M, Ford Mustang, and Toyota Celica. It can be seen from these line drawings that the chunking of these curves produces a different representation for each vehicle. Without isolating the curves in this way, the differences between these four coupes would not be as obvious. Results like these can be taken into consideration by the designer when creating new concepts. Comp 3, another 8.2% explained, is composed of the whole headlight, the inner hood line and the width of the greenhouse, as seen in Fig. 3.

Table 8 Principal component of SUVs front

	81.8%	SUV front Y	84.6%	SUV front Z
COMP 1	75.4%	Shoulder, width of body, width of greenhouse, fender, outer cowl	67.0%	Width of greenhouse, top of roof, top of headlight, cowl, top of front windshield
COMP 2	6.4%	Lower horizontal line of front bumper, outer hood line, width of greenhouse, width of body, outside of roof, outside of grill, outside of headlight, track, wheel size	11.6%	Front bumper, bottom of grill
COMP 3			6.0%	Bottom of front bumper, bottom and outside and top of headlight, bottom of grill

Table 9 Principal components of SUVs side

	81.2%	SUV side X	76.4%	SUV side Z
COMP 1	59.0%	Back of trunk, taillight, back and bottom and horizontal lines of rear bumper, back and front of rear wheel well	60.2%	Top of side window, beltline, top of roof, top of headlight, cowl, top of front windshield, top of rear windshield
COMP 2	12.0%	Top of front windshield, roof, cowl, front of door, b-pillar, beltline, top of driver's window	9.0%	Front bumper, bottom of grill, rear window, back of trunk, front of door
COMP 3	10.2%	Top and front of front bumper, grill, headlight	7.2%	Back of trunk, top and back and upper horizontal line of rear bumper, bottom of door, rocker, front and bottom and lower horizontal line of front bumper, top and bottom of tail light

Table 10 Principal components of SUVs rear

	78.1%	SUV rear Y	74.2%	SUV rear Z
COMP 1	68.3%	Fender, width of body, beltline	59.4%	Beltline, top of roof, top of side window, top of rear windshield
COMP 2	9.8%	Horizontal lines of rear bumper, inside and bottom and top of tail light, width of greenhouse, outside of roof	8.6%	Horizontal lines and top of rear bumper, top and inside and bottom of tail light, width of greenhouse, wheel size
COMP 3			6.2%	Rear bumper, width of body, outside of rear windshield, top and outside of tail light

Figure 5 shows the principal components for coupes front view vertical coordinate. This is the largest number of principal components of all vehicles, views, and coordinates, in that Kaiser's Criterion was not met until after five principal components. Since principal components 2–5 have such a low percentage, it means that the summarization of this view and coordinate is not easily simplified. Comp 1 explains 41.0% of the variance in this view and coordinate through the top of the vehicle, the top of the windshield, the cowl, the top and bottom of the grill, and the horizontal front bumper line. Many of the significant characteristics are quite intuitive and, as will be shown later, are repeated across all the vehicle classes. An additional 11.8% is explained in Comp 2 through the shoulder and fender, the wheel size, most of the headlight, and the front bumper's horizontal accent line. It should be noted that the headlight and the shoulder appeared also in the principal components for the same view horizontal axis. Their repetition here indicates that they, with their related characteristics, are important fundamental curves for coupes. Comp 3 is composed of the top, outside, and bottom of the headlight, the bottom of the grill, the horizontal

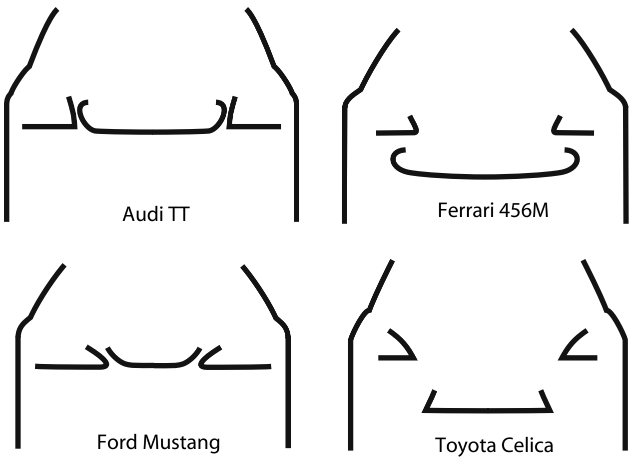

Fig. 4 Comparison of coupe front Y-axis Comp 2

front bumper accent line, the fenders, the top of the windshield and the top of the roof. Comp 4 repeats some of the previous curves, in a new grouping, and adds in the bottom of the front bumper. Finally, Comp 5 gathers some of the previous curves, brings in the whole headlight, and adds in the greenhouse. The significance of the greenhouse is that it also ties the two coordinate directions together for this view.

Figure 6 shows the principal components for coupe side view x-coordinate, which are summarized in Table 3. In this instance, there are only 2 principal components that fall within Kaiser's Criterion. Surprisingly, over half (59.6%) of the differences between coupes in this view and coordinate is explained by the rear and bottom of the trunk, half of the taillight, the top of the front windshield, the top and bottom of the rear bumper, and just part of the rear wheel well. This chunk is not at all intuitive and offers some insight into the differences between coupes. Comp 2, 17.5%, is composed of nose curves: the headlight, the front bumper, grill, the front of the hood, and the front edge of the front wheel well. Since this chunk is focused on one area of the vehicle, a vehicle designer can be certain that choices made in the form of the nose will have a strong influence on the design of the vehicle when viewed from the side.

The first principal component for coupe side view z-coordinate (Fig. 7) ties together the top of the grill, the bottom of the windshields, the top and bottom of the side window, and the top of the roof, to explain about a third of the variation between coupe in this view and coordinate. The relationship between these curves is somewhat obvious, but can be utilized by a vehicle designer. Comp 2 is split between the horizontal front bumper line, the taillight, the horizontal crease, top, and bottom of the rear bumper, and the top and outside curves of the trunk. Again, it is not

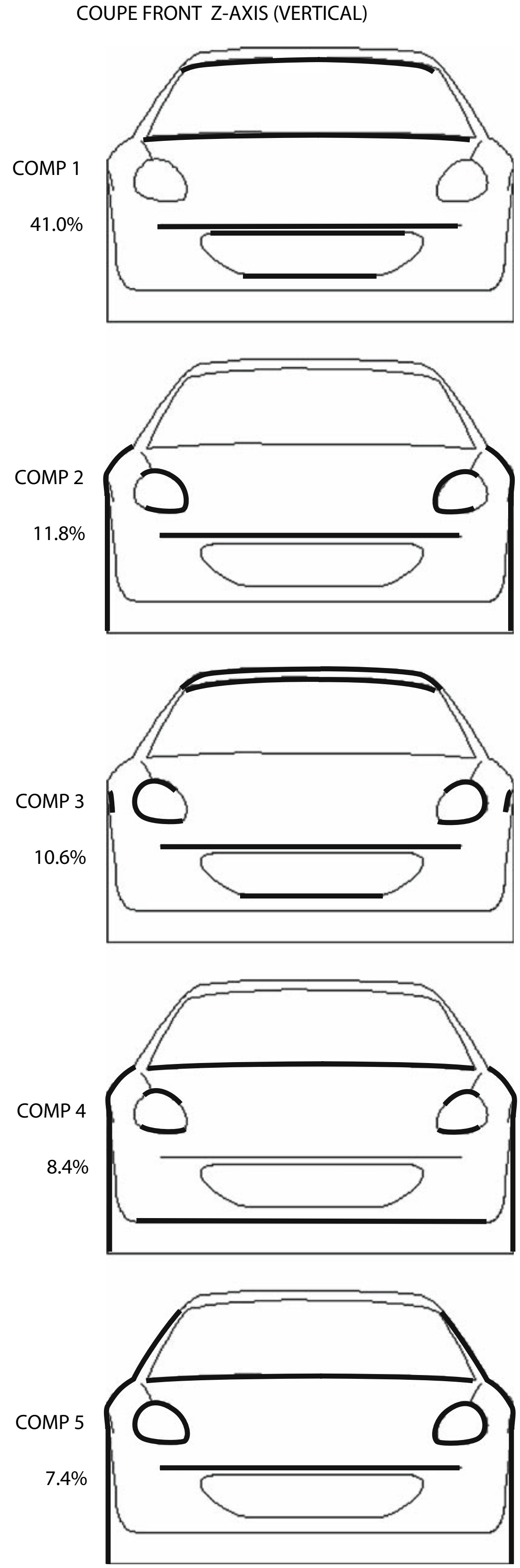

Fig. 5 Coupe front Z-axis principal components

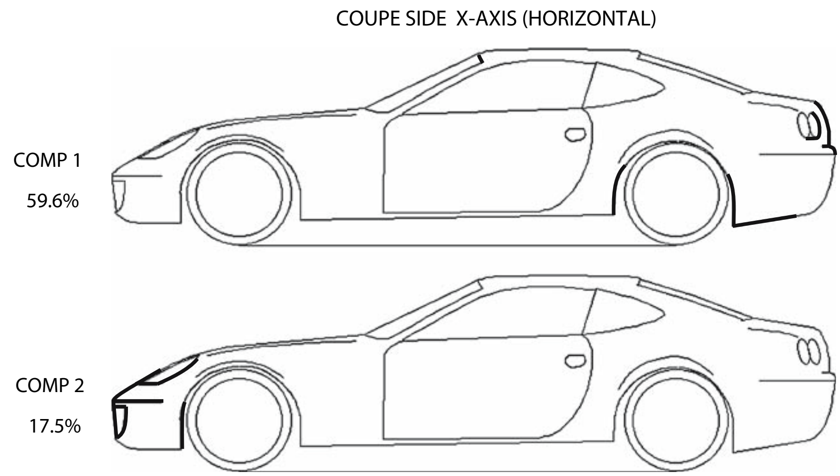

Fig. 6 Coupe side *X*-axis principal components

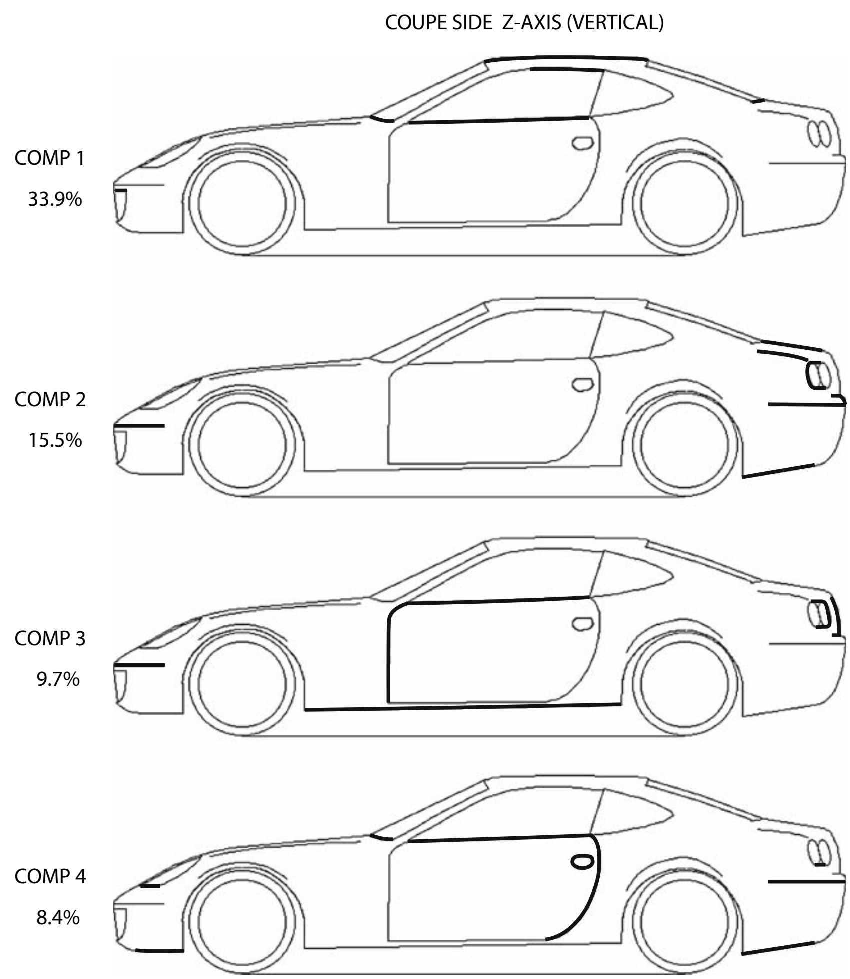

Fig. 7 Coupe side *Z*-axis principal components

obvious that these very distant curves would chunk together to account for 15.5% of the differences between coupes in this view and coordinate. In Comp 3, the horizontal front bumper line, part of the taillight and the rear of the trunk are brought together. The addition of the rocker and the front and top edges of the door ties this chunk to Comp 1 of this view. Finally in Comp 4, which accounts for 8.4% of the variation between coupes, the chunk repeats the cowl, the top of the side door, and parts of the rear bumper and taillight. The addition of the rear of the

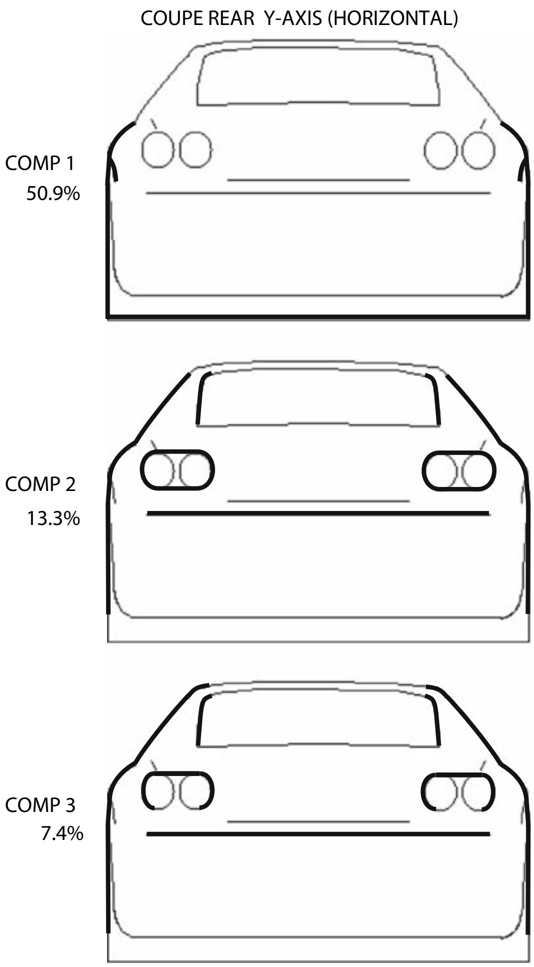

Fig. 8 Coupe rear *Y*-axis principal components

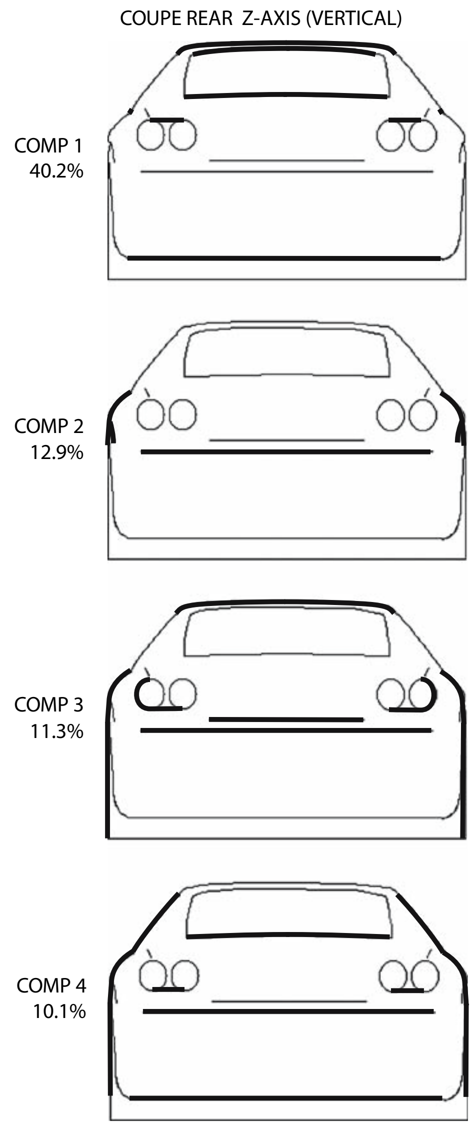

Fig. 9 Coupe rear *Z*-axis principal components

door, the door handle, the bottoms of the headlight and front bumper form connections to other principal components in the side view.

Figure 8 shows the composition of the principal components for the rear view, *y*-coordinate of coupes which is detailed in Table 4. Comp 1 explains half (50.9%) of the variation through the width of the body, the fenders, the track width and the wheel size. This is quite similar to Comp 1 in the front view, which is understandable. The consistency between these two views further verifies this method of chunking together fundamental shapes through principal component analysis. Comp 2 chunks together the width of the body and greenhouse with details that are rear view specific: the taillights, the horizontal rear bumper crease, and the outside of the rear windshield. This, again, echoes the principal components in the front view. Comp 3 explains an additional 7.4% by grouping together the curves from Comp 2, minus the bottom of the taillight and adding the outside of the roof. The relationship between the chunks in the front and rear view will enable the designer to address these curves together.

Figure 9 shows the four principal components for the rear of coupes in the vertical direction. The first principal component chunks together the top of the roof, the top and bottom of the rear windshield, the top of the taillight, the bottom of the rear bumper, and the belt line for an explanation of 40.2% of the variation. This is similar to the chunking for Comp 1 in the front view, vertical direction. Comp 2 includes the shoulder and the horizontal bumper line, like the front view. It also draws in the fender. Comp 3, for 11.3%, combines the roof, the trunk, the horizontal bumper crease, the shoulder, the outside and bottom of the taillight, and the wheel size. Finally, Comp 4 combines the previous characteristics in a new chunk with the bottom of the rear windshield, the bottom of the taillight, the horizontal crease and bottom of the rear bumper, and the

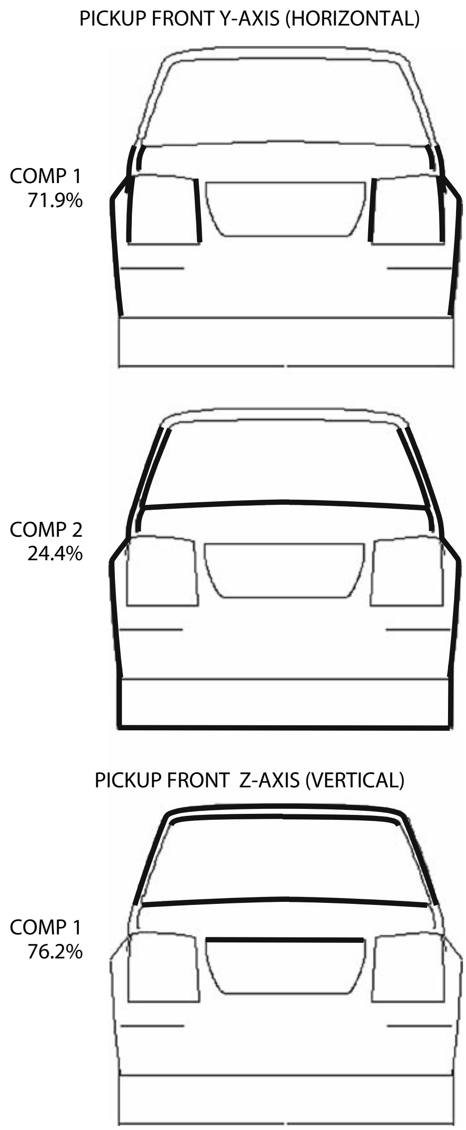

Fig. 10 Pickup front principal components

shoulder. It also adds in the width of the body and greenhouse, to tie it into the principal components for the horizontal direction.

These vehicle characteristics, along with the rest of the significant characteristics from the other views and coordinates, are the fundamental characteristics within the coupe vehicle class. It is interesting to note that while some of the chunking of the characteristics is quite intuitive (i.e. the width of the vehicle in the front and rear views), some of them are non-obvious (i.e. the bottom and outside of the grill and inside and bottom of the headlight in the front view). Many of the characteristics that stand out in a certain view are significant in both the horizontal and vertical directions. This indicates that these characteristics are fundamental between the vehicles in the vehicle class.

4.2 Results for pickups

Tables 5, 6 and 7 summarize the principal components for pickups, and are shown visually in Figs. 10, 11 and 12. Due to the small sample size and the gross similarities between pickups, Kaiser's Criterion eliminated all but the first principal component for each view and coordinate, except for front view *y*-coordinate, which also has a second principal component. Comp 1 (Fig. 10) for the front view y-coordinate pulls in many of the same curves used in the principal components for coupes: the width of the body, the outer hood line, the inside and outside of the headlights, and the fender. Comp 2 chunks together the rest of the characteristics: the width of the body and greenhouse, the cowl, the outside of the front windshield, the track width, the outer hood line and the wheel size. These two principal components account for an astonishing 96.3% of the variation in pickups in this view and direction. It is interesting to note that neither the grill nor the front bumper appears in either chunk. The *z*-coordinate is similar as that for the coupe Comp1: the height of the roof, the top of the windshield, the cowl, and the top of the grill. The greenhouse is an addition and helps to connect the two directions in this view. The side view *x*-coordinate (Fig. 11) chunks together the front of the vehicle, the top of the roof, and the back of the door. It also, uniquely, includes the rim size and the door handle, two characteristics that would seemingly not have such an effect on a vehicle's differentiation. The side view vertical principal component has many of the same curves as the coupes. The top of the vehicle, the cowl, the top and back of the bed (which is the trunk in a coupe). The door handle appears here, again, tying this coordinate to the horizontal direction. The fenders are a new addition. This chunk is quite telling in the way the curves around the beltline are grouped. Additionally, the door handle's second appearance confirms that this characteristic, while often an afterthought in the design process, is a strong differentiator between pickups. As a counter point, the wheel wells, which are often considered drivers for brand differentiation in pickups, are not significant. For example, Ford trucks maintain a constant arc wheel well and Chevrolet uses a rectangle with rounded corners. While they may be important according to the brand definition, alone they are not enough to differentiate between other pickups. This may be due to the fact that other pickups also share similarly shaped wheel wells. Many of the subtle details in form of the headlights were left off, which also help to distinguish pickup trucks. The fact that the general forms for these characteristics do not vary much between the pickups provides designers with a direction to head if they wish to strongly differentiate their vehicle. For example, round or diamond-shaped headlights would stand out strongly among pickups.

Fig. 11 Pickup side principal components

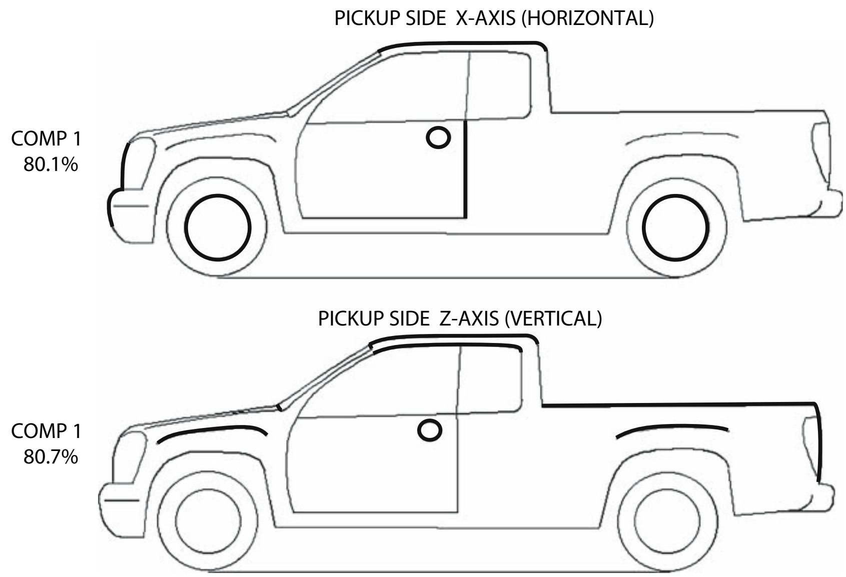

Figure 12 shows the principal components for the rear of pickups. Comp 1 in the *y*-coordinate accounts for 93.8% of the differentiation between pickups. The width of the greenhouse and body, the fender, and the outside and top of the taillight are not surprisingly the only characteristics that can be modified due to the strong functional considerations in the design of the rear of pickups. The vertical direction is explained through the top of the roof, the bottom of the rear windshield, to top of the bed, and the greenhouse. Again, many of the characteristics are not fundamental due to their functional requirements, i.e. the bumper height and tail gate width.

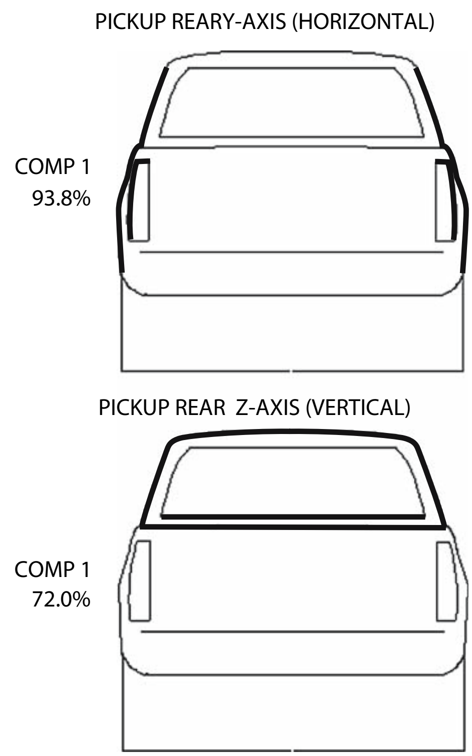

Fig. 12 Pickup rear principal components

4.3 Results for SUVs

Figure 13 shows the principal components for the front of SUVs in the horizontal direction, which are listed in Table 8. The first principal component in each direction dominates the differentiations in the class. In the horizontal direction, the width of the greenhouse and the body, and the outer cowl account for 75.4% of the variation among the SUVs. An additional 6.4% is accounted for through the chunking of the outside of the grill and headlight with the outer hood line and the width of the greenhouse and body, the outside of the roof, the track width and wheel size in Comp 2. It is important to note that the width of the greenhouse and body shows up in both of these principal components. This indicates that the greenhouse and body width is important to differentiating between SUVs, regardless of which other characteristics it is chunked with.

The SUV front vertical principal components (Fig. 14) again have similar characteristics to the coupes and pickups, but are chunked slightly differently. 67.0% of the variation is explained through the top of the vehicle, the top of the windshield, the height of the cowl, and the top of the headlights. Comp 2, an additional 11.6%, chunks the bottom of the grill, the bottom of the front bumper, and the horizontal crease in the front bumper. Finally, 6.0% of the vertical variation in the front view of SUVs is explained through the bottom of the front bumper and grill, and all but the inside of the headlight.

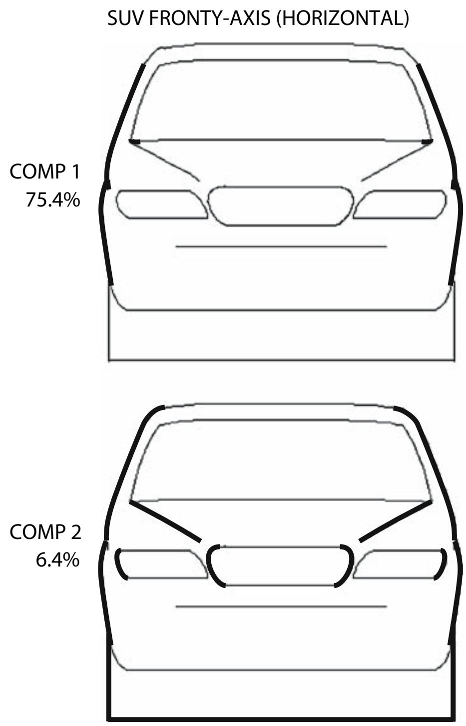

Fig. 13 SUV front *Y*-axis principal components

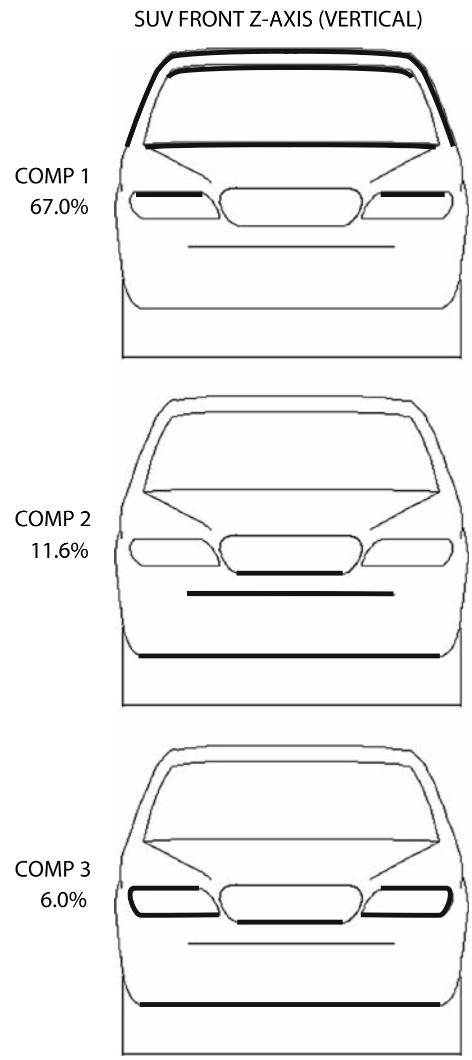

Fig. 14 SUV front *Z*-axis principal components

Figure 15 shows the principal components for the SUV side view *x*-coordinate, which are described in Table 9. Comp 1 is similar in composition to coupes in that it focuses on the rear of the vehicle. Here, the taillight, the whole of the trunk, rear bumper and rear wheel well tell of the differences between SUVs. Unlike the other vehicle classes, Comp 2 at 12.0% includes the roof, top of the windshield, the cowl, B-pillar, the front and top of the driver's door, and the belt line. Neither the coupes nor the pickups have so much of the greenhouse in any of their side view horizontal principal components. 10.2% of the variation between SUVs in this view and direction is explained through a grouping of the grill, front bumper, and headlight. This is similar to Comp 2 for this view and coordinates for coupes.

The roof and top of the side windows again dominate the difference between vehicles in this class, as shown in the first principal component for the side view vertical direction (Fig. 16). These, with the belt line, the cowl, and the top of the headlight, describe over half (60.2%) of the difference between SUVs. Comp 2, while significantly less at 9.0%, brings in front of the door with the front bumper, the rear side window, and the back of the trunk. This second appearance of the front of the door indicates that shut lines are important. Often, manufacturing desires that these lines be as subtle as possible. But, these can be used to further differentiate between SUVs. Curiously, Comp 3 is quite spread out. It is composed of the tip of the front and rear bumper, the back of the trunk, the top and bottom of the taillight, the rocker and the bottom of the doors. This would indicate that the vertical height of the chassis (ground clearance) is an important characteristic.

Figures 17 and 18 show the principal components for the rear of SUVs, which are listed in Table 10. In the horizontal direction, 68.3% of the variation is determined by the width of the body, shown in Comp 1. An additional 9.8% in Comp 2 is through a chunking of the details of the rear bumper, the outside of the roof and greenhouse, and all but the outside of the taillight. In the vertical direction, over half (59.4%) of the variation is described through the chunking of the top of the roof and the rear windshield in Comp 1. It is important for designers to consider that the interaction of just three curves (the top of the roof, the outside of the roof, and the top of the rear windshield) is so fundamental to a design when viewed from the rear. Comp

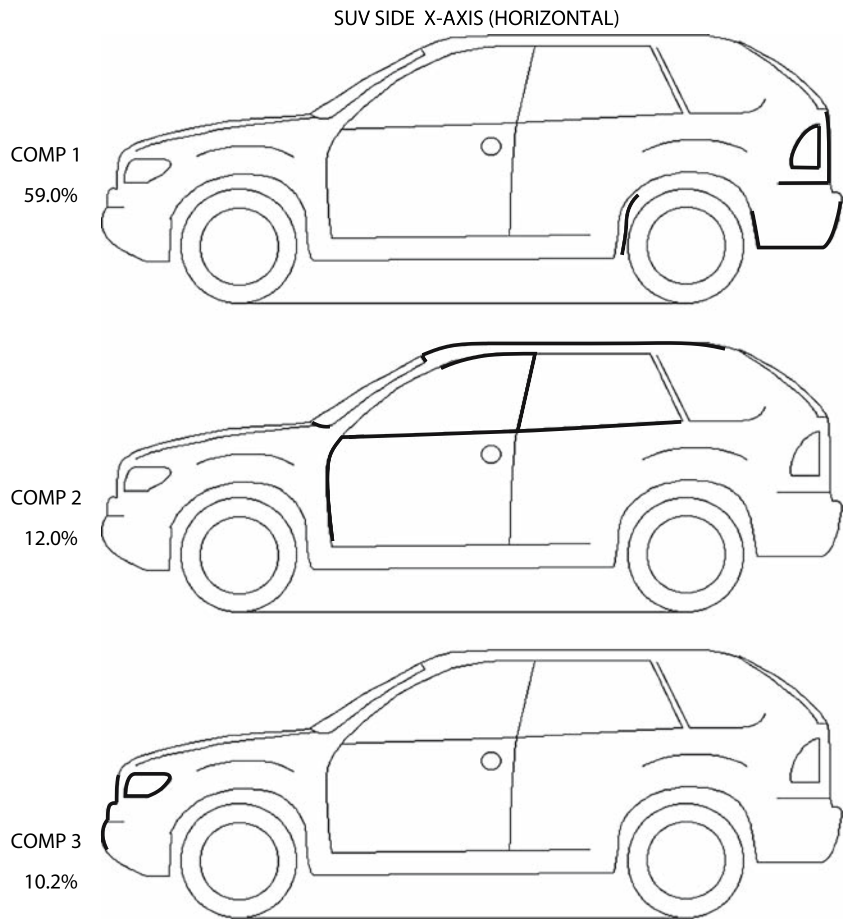

Fig. 15 SUV side *X*-axis principal components

2 brings in the characteristics of Comp 2 in the horizontal direction, minus the outside of the roof, with the wheel size. Finally, Comp 3 is composed of the entire rear bumper, the width of the body, the outside of the windshield, and the rest of the taillight. It is interesting to note that only in coupes did the track width stand out as a differentiation in the rear view.

While this analysis determined which curves chunked together are the fundamental characteristics for vehicles in a particular class, it also showed us the similarities. If Figs. 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 and 18 are revisited and the curves that are not highlighted are focused on, another story evolves. These are the curves that are parametrically consistent across all the vehicles within that class. In every class the hood, the front windshield, the A-pillar, and the rear windshield were never indicated as differentiators. The similarities may be due to manufacturing or engineering constraints. These, and other class specific similarities, can be used by the designer to truly differentiate their vehicle. One way to differentiate would be to take these characteristics and purposefully create them in a way that matches a different vehicle class.

5 Application

Vehicles can now be created based upon an understanding of the shape relationships derived from the principal component analysis. Figures 19 and 20 show two vehicles, side view only. Figure 19 is a coupe concept created by drawing the dominant principal components first and then filling in the rest of the vehicle (Figs. 6, 7). The curves in Fig. 19a were chosen based upon the percentage of variance explained: Comp 1 *X*, Comp 1 *Z*, Comp 2 *X*, Comp 2 *Z*, Comp 3 *Z*, then Comp 4 *Z*. The parameters for the curves were chosen with the intention to maximize the greenhouse and flatten the belt line (contrary to the vehicle design push in 2006). Because the curves that are chunked through the principal component are drawn first, the significant characteristics are addressed first. The rest of the vehicle is then filled in to complete the vehicle in Fig. 19b. While this vehicle may seem cartoonish, it shows that by addressing the principal component chunks first, a unique vehicle concept can easily be created.

As a contrast, Fig. 20 is a coupe concept created by focusing on the curves not highlighted in the principal components. This fanciful vehicle arises due to the fact that

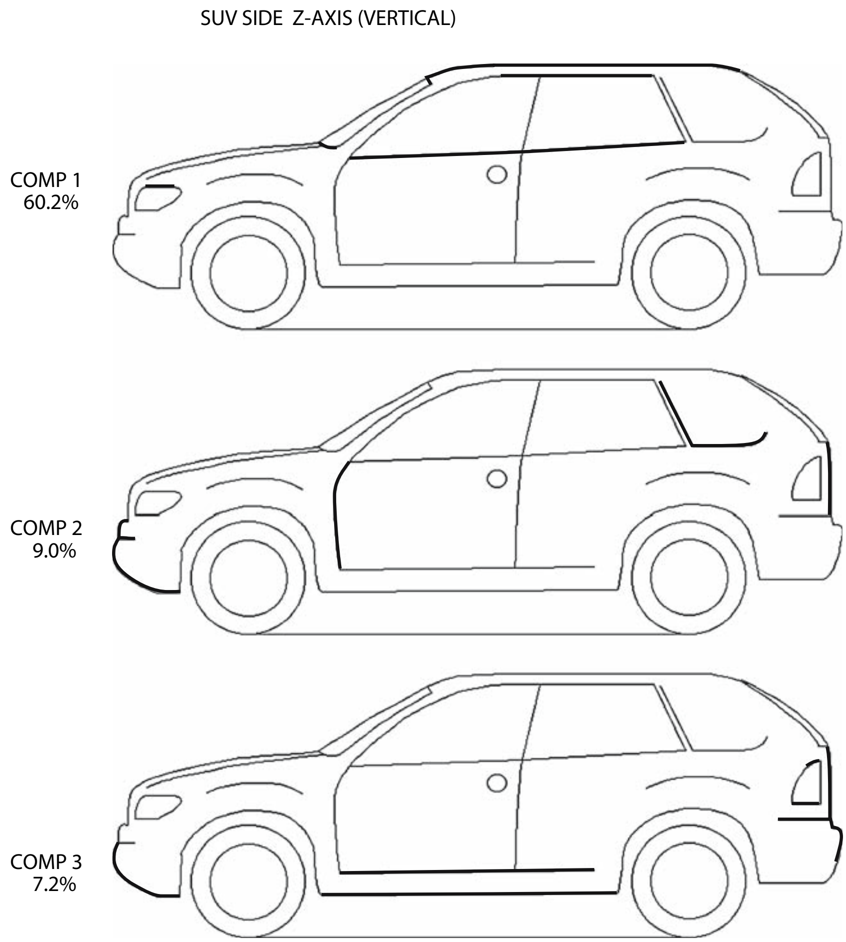

Fig. 16 SUV side *Z*-axis principal components

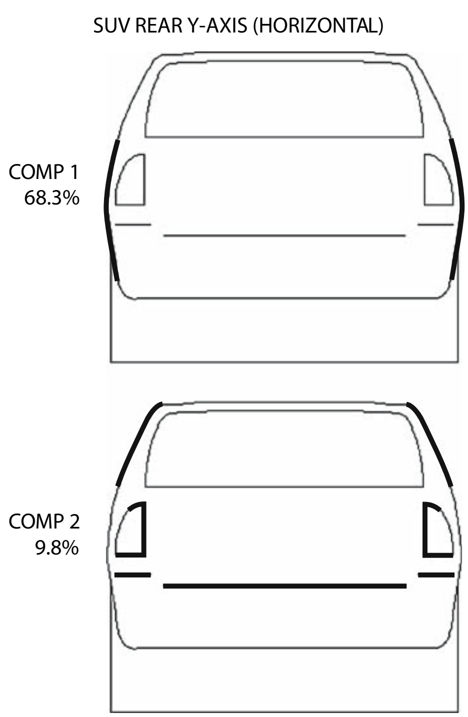

Fig. 17 SUV rear *Y*-axis principal components

the curves chunked into principal components are purposely ignored until after the design is started, as seen in Fig. 20a. By directly addressing the curve chunks, or purposely ignoring them, a vehicle designer can push the initial concept in a unique direction. The designer no longer needs to start from a specific feature and move forward. The designer can start with the curves ascertained to be important (higher percentage of variance explained) by the principal component analysis and then expand the design from these fundamental curves. By approaching the concept creation processes from this unique point of view, the vehicles are pushed into novel forms.

These two vehicles are a glimpse into the potential of new vehicle concepts that could be created. The curve chunking also provides the opportunity for a formalized curve generation methodology, like a shape grammar. The chunks in the principal components, automatically derived, could be the vocabulary used to build shape grammar rules (Stiny 1980).

6 Conclusions

Vehicles have many different characteristics, i.e. doors, wheels, hood, etc. Each of these characteristics can be

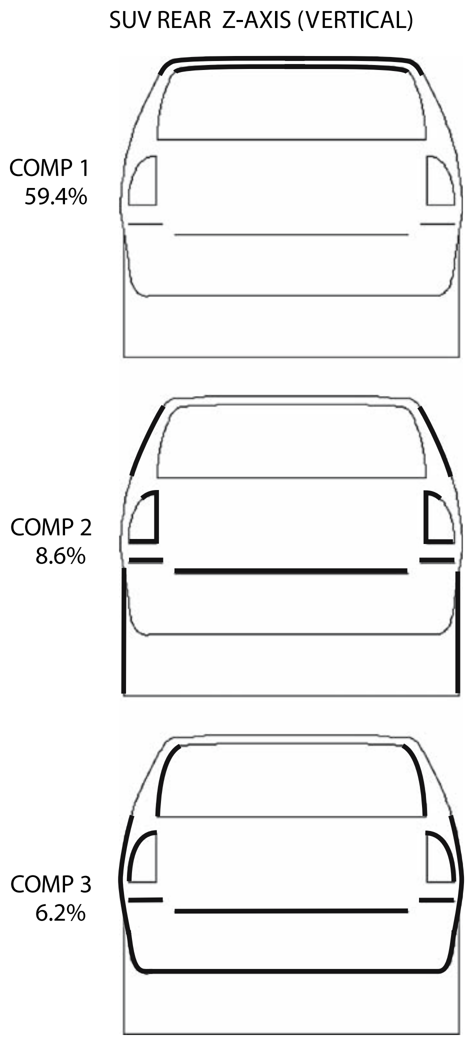

Fig. 18 SUV rear Z-axis principal components

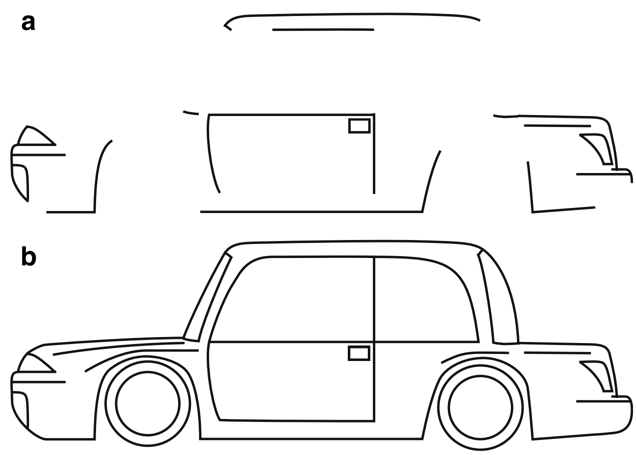

Fig. 19 Coupe concept by applying principal components

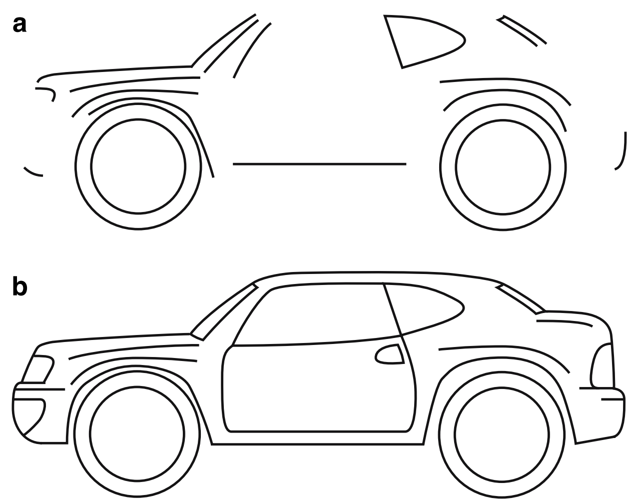

Fig. 20 Coupe concept by ignoring principal components

represented with one or more curves. Traditionally, as a vehicle is designed, the spatially related characteristics influence each other the most. They are often created in sequence: the wheels effect the wheel wells which effect the fenders and so on due to proximity. By statistically analyzing a class of already established vehicles, new vehicle characteristic relationships can be determined based on their description of the vehicle, not on proximity. In this paper principal component analysis was used to find the fundamental vehicle characteristics within the vehicle classes of coupes, pickups, and SUVs. The analysis chunked together specific curves and assigned a percentage variance explained to each chunk. These chunks compose the foundation of the vehicle design. A vehicle designer can take these foundational chunks into consideration when designing new concepts, with the full knowledge that the relationship between chunked curves will have a strong effect upon the gestalt of the final design of the vehicle. The methodology introduced in this paper is applicable to any product class that is composed of physical characteristics shared across the product class.

Acknowledgments Funding for this research was partially provided by the National Science Foundation under grant DMI-0245218 and General Motors.

References

Biederman I (1987) Recognition-by-components: a theory of human image understanding. Psychol Rev 94:115–147

Csere C (2003) Infinity FX35. Car Driver 49:140

Kaiser HF, Rice J (1974) Little Jiffy. Mark IV. Educ Psychol Meas 34:111–117

Krzanowski WJ (1988) Principles of multivariate analysis: a user's perspective. Oxford University Press, Oxford

McCormack J, Cagan J (2002) Supporting designer's hierarchies through parametric shape recognition. Environ Plann B Plann Des 29:913–931

Orsborn S, Cagan J, Pawlicki R, Smith R (2006) Creating cross-over vehicles: defining and combining vehicle classes using shape grammars. Artif Intell Eng Des Anal Manuf 20:217–246

Patel N, Shmueli G, Bruce P (2006) Data mining in Excel. Resampling Stats, Inc., Arlington

Prats M, Earl C, Garner S, Jowers I (2006) Shape exploration of designs in a style: toward generation of product designs. Artif Intell Eng Des Anal Manuf 20:201–215

Stiny G (1980) Introduction to shape and shape grammars. Environ Plann B 7:343–351