A QUANTITATIVE MODEL FOR IDENTIFYING REGIONS OF DESIGN VISUAL ATTRACTION AND APPLICATION TO AUTOMOBILE STYLING

Y. Pan, A. Burnap, Y. Liu, H. Lee, R. Gonzalez and P. Y. Papalambros

1. Introduction

The aesthetic appeal of designed artifacts has been long recognized as significantly affecting customer preferences; examples include golden section proportions, Gestalt Psychology, form versus function, emotional design and craftsmanship. Both practicing designers and design researchers have focused on this important topic, see, e.g., [Coates 2003], [Norman 2004], [Osborn 2009], [Reid 2010], [Reid 2013]. This notion is particularly true for increasingly commoditized products such as automobiles, as standardization across product components and manufacturing processes are pushing product differentiation to moreso to perceptual attributes such as aesthetic styling and corresponing visual attraction [Bloch 1995], [Moulson 2008].

To better understand the factors affecting visual attraction, we extend previous research on both descriptive and predictive aspects of aesthetic appeal. Descriptive studies of aesthetic appeal have examined the saliency of design features and their propensity to draw perceptual attention [Crilly 2004]. Berlyne's theory of appeal, for example, aggregates sensory information and models aesthetic appeal by balancing novelty and arousal and trading off meaning and recognition [Berlyne 1971]. The designer's intent is often focused on actively drawing visual attention to salient regions of a design in a communication between designers and customers [Mono 1997], [Crilly 2004]. This communication may begin with a pleasant initial impression of the customers due to the attractive appearance of the design and cement that impression by expressing attributes important to them [Norman 2004].

Predictive studies of factors that affect aesthetic appeal model which design features evoke particular visual design attributes. Linear models of forward communication such as conjoint analysis [Ben-Akiva 1999] and Kansei engineering [Nagamachi 1995] have been used to capture and predict design attributes as functions of design features. These models may use design features implicitly 'learned' [Orsborn 2009], [Ren 2013], hand-crafted features [Orsborn 2009], [Petiot 2010], [Reid 2010], [Kelly 2011], or learned through dimensionality reduction [Yumer 2015]. Another approach is to use eye-tracking methods where the subjects' gaze and fixation time to a given design stimuli are measured and correlated to behavioural information such as consumer choice [Reid 2013], [Du 2014], [Marshall 2014].

To quantiatively capture these descriptive and predictive factors of aesthetic appeal and visual attention, we adopt the framework of design as a communication between designers and customers. This framework suggests that design communication occurs from designer to customer, hereafter referred to as forward communication. We extend this framework to include communcation from customer to designer, or a backward communication direction of customer response. This forward- backward design

communication concept shown in Figure 1 draws heavily from previous literature but the formalism introduced in this paper is novel.
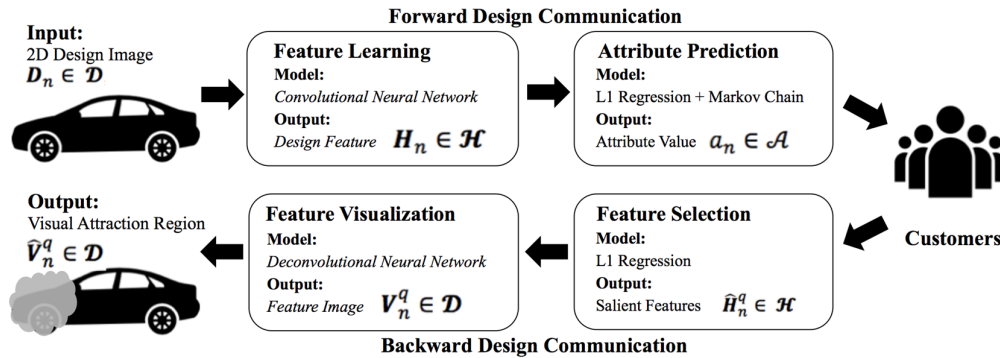
Figure 1. Overview of design process using the proposed quantitative communication model
The goal is to predict a region of visual attraction, denoted in grey given a particular design

In the forward communication direction, the high dimensionality of realistic design representations and complexity of the nonlinear mapping between this representation and an attribute value creates a challenging statistical estimation problem [Burnap 2015]. Sophisticated nonlinear models have been developed to model this process with high accuracy, such as kernel methods [Ren 2013] and feature learning [Pan 2015]. With nonlinear models, predictive performance of the underlying physics is significantly improved at the cost of reduced interpretability. With linear models, interpretability is often possible but predictive power is relatively poor due to assumptions that typically do not hold, such as linearity, feature independence, homogeneity, and complex noise distributions.

In the backward communication direction, inverting a nonlinear function to model backward communication poses significant challenges. The backward process is often quantified using experiment-based approaches such as eye-tracking and stated responses [Duchowski 2002], [Chang 2013], [Du 2014], [Marshall 2014]. These approaches work well in analysing overall aesthetic performance, but do not typically provide information about each aesthetic attribute separately. Moreover, these backward approaches do not currently use information from the forward communication.

Motivated by a collaboration with practicing automotive designers, our research goal is to capture this forward and backward communication by identifying regions of a design that draw visual attention. We introduce a data-driven method to simultaneously quantify both the forward and the backward communication. This method does not require humans to directly provide attention data, instead this method predicts the attention region in the given design from humans' feedback on its attribute values in four stages: (i) feature learning, (ii) attribute prediction, (iii) feature selection, and (iv) feature visualization. The resulting mathematical model has three goals: (i) assess aesthetic attributes based on the design representation (in our application study these are pixel-based 2D images); (ii) invert the nonlinear model to predict corresponding attention region; and (iii) leverage useful information from both communication directions.

The modeling tools employed consist of a convolutional neural network, L1 regression, a crowdsourced ranking Markov chain, and a deconvolutional neural network. The four data sources we use for modeling are summarized in Table 1. We conducted an experiment involving four steps: (i) learn design features of 2D car images through a convolutional neural network trained by ImageNet [Deng 2009] and Flickr [Karayev 2013] data sets; (ii) use L1 regression to model the relation between the design features and the design attribute values determined by a crowdsourced ranking Markov chain; (iii) determine salient features according to the L1 regression model; and (iv) determine visual attraction regions by visualizing the selected salient features using a deconvolutional neural network. The L1 regression was chosen to introduce sparcity thus reducing the complexity of the number of design features needed to relate to the design attributes. The major contribution of this research is the the extension of previous quantification of *forward only* design-customer communication to a combined *forward-backward* communication using a purely data-driven approach and multiple large scale data resources. The purely data-driven

INDUSTRIAL DESIGN

approach can also be used alongside existing methods such as eye tracking and dimensionality reduction of 2D and 3D designs.

Table 1. Description of the four data sets used in this work

Dataset	ImageNet	Flickr	Vehicle Images	Design Attribute Ranking
Num. of Data	15,000,000+	80,000	110	5,054
Annotation	Object Name	Style Annotation	Vehicle Make an Model	Design Attribute
Source	Open Source	Open Source	Image Search Engine	Crowdsourcing

2. Related work

We build on literature from visual attention studies from art and product design, and data features for representing 2D images from biology, computer vision, and the design community.

2.1 Visual attention in design

Perhaps the earliest experimentally recorded investigation of design attention was conducted to analyse regions of eye-gaze fixation of 55 artistic pictures by 200 participants [Buswell 1935]. Such eye-tracking approaches have been successful in optimizing the layout of product placement, advertisements, and labelling objects in supermarkets. Readers are referred to [Duchowski 2002] for a comprehensive and crossdisciplinary review of eye-tracking research. Recently, these methods have been applied to design research, including vehicle face attribute assessment with Kansei engineering [Chang 2013], design representation comparison [Reid 2013], relations with vehicle face component size changes [Du 2014], and technical diagram assessment [Ruckpaul 2015].

2.2 Data features for design representation

We review data features from several perspectives: biology, computer vision, and design. A feature is a general term for a function of the underlying design variables, used to represent the design at a particular level of fidelity. For example, complex 3D meshes underlying a realistic design concept can also be represented by a set of control points. This set of control points may be a more efficient the feature representation of the realistic design concept, as it is able to preserve the important design information in the space with lower dimensions [Ren 2013], [Yumer 2015].

At a neurophysiological level, visual attention can be modeled in a *bottom-up* fashion according to perception pathways [Hubel 1962]. Such pathways are analogous to the *forward* direction of our model, from the 2D design image space to the design attribute space, see Figure 1. Similarities have been shown between neural network data features and Gabor features [Marčelja 1980] known to model visual cortex V1 and V2 cell receptive fields [Lee 2008].

There is vast amount of foundational and ongoing work from computer vision researchers on handcrafted image features and implicitly-learned (i.e., learned purely from data) image features. Handcrafted features tend to outperform implicitly-learned features due to the reduction in the uncertainty of the true data-generating mechanism. For example, features learned for face recognition take advantage of facial symmetry and facts such as two eyes are separated by a nose and mouth. One the other hand, implicitly-learned features in so-called "feature extraction" tend to be more general for a variety of tasks. Such features include HOG features [Dalal 2005], and features learned in convolutional neural network [Krizhevsky 2012].

There are data features specific to design, for example, in investigating how design attributes vary according to corresponding variability in a design representation. These design representations may be hand-crafted, such as a set of parametric handles to manipulate vehicle silhouettes [Petiot 2010], [Reid 2010], [Poirson 2013]. These design representations have also been created implicitly using finite shape grammars that together form more complex representations [Pugliese 2002], [McCormack 2004], [Orsborn 2006]. Hybrid approaches that learn the set of handles have been studied, for example, autoencoders for 3D object manipulation to affect attribute ratings [Yumer 2015], and representations that combine hand-crafted and implicitly learned representations to capture design freedom and brand recognition [Burnap 2015].
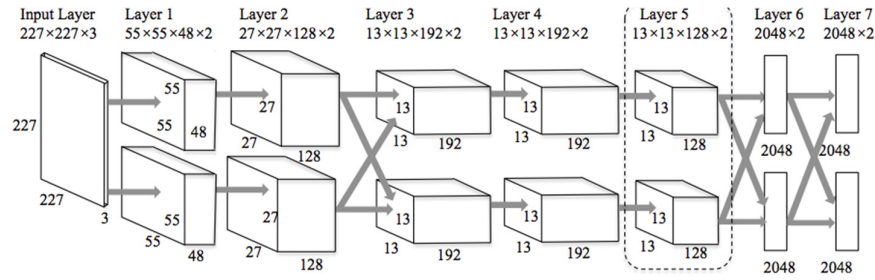
Figure 2. AlexNet convolutional neural network structure

3. Method

We model how customers perceive aesthetic design attributes and build a mapping from the design image space $\mathcal{D}$ to the attribute space $\mathcal{A}$ through an intermediate step in the design feature space $\mathcal{H}$, and then inverting this design-attribute mapping to predict visual attraction regions $\mathcal{V}$ in the original design image space $\mathcal{D}$. The four modeling steps--feature learning, attribute prediction, feature selection, and feature visualization--are detailed below as well as in Figures 2–4.

3.1 Feature learning using deep convolutional neural network

A deep convolutional neural network is a hierarchical model consisting of multiple layers (which could be conceptualized as layers of neurons following the organization of neurons in the human cortex), with each layer extracting higher-level data features from the previous layer. The output of each layer is a collection of features of the input image. Recent research has successfully applied these deep convolutional neural network features to aesthetic related tasks such as style recognition [Karayev 2013] and artistic image generation [Gatys 2015].

The features learned from a deep convolutional neural network depend on its structure and training data. Here, we learn design features using the structure of AlexNet [Krizhevsky 2012], detailed in Figure 2, due to its record-beating performance on the ImageNet 2012 classification benchmark [Deng 2009]. Originally, AlexNet was trained on the ImageNet dataset, which consists of over 15 million images with over 22,000 class labels of the objects in image (e.g., dog breeds and strawberries). In addition, further fine-tuning was obtained by using additional images from the Flickr dataset [Karayev 2013], which itself consists of 80,000 images with more-specific style labels (e.g., 'melancholy,' 'ethereal') to modify higher layers to be more specific to desired aesthetic concepts. [Karayev 2013] show that mid-level features (layer 5 and layer 6) in AlexNet outperforms hand-tuned features in style recognition tasks and achieves the same level of prediction accuracy as participants in Amazon Mechanical Turk in a photographer group membership prediction task.

Accordingly, for any design image $D_n \in \mathcal{D}$, using the deep convolutional neural network with structure and training procedure described above, we choose the feature outputs in layer 5 (see Figure 2) as the feature representation of the input design image, as these design features contain both the 2D image-specific information (usually contained in lower layers) and design attribute-specific information (usually contained in higher layers); we denote this feature representation as $H_n \in \mathcal{H}$.

3.2 Design attribute prediction using crowdsourced Markov chain and L1 regression

3.2.1 Crowdsourced Markov chain

To obtain design attribute values for each 2D vehicle design image (e.g., the Toyota Prius may be 0.07 aggressive and 0.86 youthful), we assumed a ranked list of all 2D vehicle designs for each attribute. To obtain these ranked lists, we crowdsourced evaluations in the form of partial ranked lists partial ranking into a full ranking by an aggregation model based on Markov chain theory. Specifically, we assumed that the full ranking corresponds to the probability mass of individual designs of the stationary distribution of an ergodic Markov chain. We obtain the stationary distribution by using a modified version of the PageRank algorithm [Brin 1998]; see [Burnap 2015] for more implementation details.
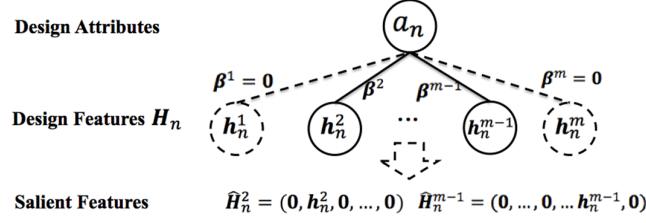
Figure 3. L1 regression

3.2.2 L1 regression

Previous research has shown that transforming highly nonlinear design variable relationships into more easily human-memory "chunked" perceptual attributes justifies the linear models commonly used in the design community [Hubel 2014]. Accordingly, we model the relation between design attribute and design features as a L1 regularized regression model. Given the design feature representation $\boldsymbol{H}_n$ for Design $\boldsymbol{D}_n$ as well as its design attribute value a_n, we assume that there is a linear relationship between a_n and $\boldsymbol{H}_n$:

$$a_n = \boldsymbol{H}_n \boldsymbol{\beta} + \epsilon_n \tag{1}$$

where ϵ_n is a Gaussian distributed random variable. To determine the coefficient vector $\boldsymbol{\beta}$, we minimizes a loss consisting of the distance between the design attribute value and its estimation as well as a L1 regularization on $\boldsymbol{\beta}$, as given in Equation (2), in which the parameter α is determined by cross validation as is common in L1 regularization methods (Equation 2)

$$\boldsymbol{\beta} = \frac{argmin}{\boldsymbol{\beta}_0} \sum_{n=1}^{N} \left\lVert a_n - \boldsymbol{H}_n \boldsymbol{\beta}_0 \right\rVert_2 + \alpha |\boldsymbol{\beta}_0|_1 \tag{2}$$

The role of the L1 regularization is to reduce the dimensionality of $\boldsymbol{\beta}$ according to the shrinkage parameter $\boldsymbol{\alpha}$.

3.3 Salient feature selection using attribute prediction model

The L1 regularized linear regression in attribute prediction estimates the coefficient vector $\boldsymbol{\beta}$, where only some of its elements are non-zero. The features corresponding to those nonzero coefficients are modeled to influence the attribute. Based on this idea, we define the salient coefficient set

$$S = \{p | \boldsymbol{\beta}^p \neq \boldsymbol{0}, \boldsymbol{\beta} = [\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, \dots, \boldsymbol{\beta}^m]\}, \tag{3}$$

where m is the number of features, and the salient feature representation for design image $\boldsymbol{D}_n$ is:

$$\widehat{\boldsymbol{H}}_n^q = \left(\boldsymbol{0}, \dots, \boldsymbol{0}, \boldsymbol{h}_n^q, \boldsymbol{0}, \dots, \boldsymbol{0}\right) \in \mathcal{H}, q \in S. \tag{4}$$

This representation contains only one influential factor of the attribute. Using this feature representation, we are then able to apply the following feature visualization method to separately visualize the influential factors.

3.4 Feature visualization using deconvolutional neural network

A deconvolutional neural network may be considered an approximated inverse mapping of a convolutional neural network [Zeiler 2014]. This inverse mapping is achieved by inverting the operations in the original convolutional neural network in the reverse sequence. In our model, the salient feature representation $\widehat{\boldsymbol{H}}_n^q$ is passed as input to the deconvolutional neural network attached to AlexNet. Successive layers are inverted until we reach the input pixel space. This operation allows us to obtain a feature image $\boldsymbol{V}_n^q$ in the design image space $\mathcal{D}$, which contains only the pixel information
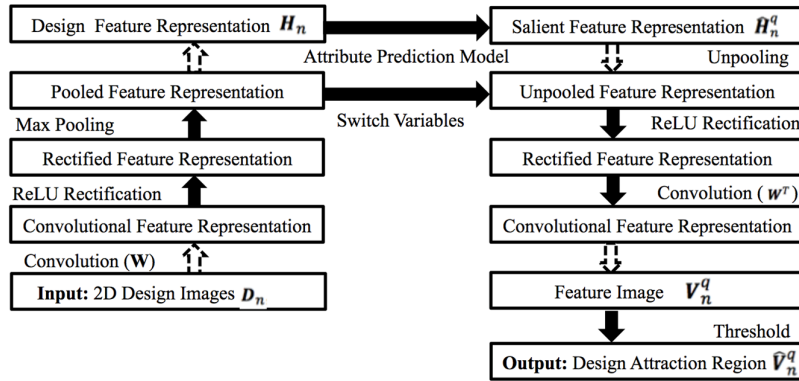
Figure 4. Deconvolutional neural network method flow

that influence the salient feature, which itself most influences the desired design attribute. The attraction region $\widehat{V}_n^q \in \mathcal{D}$ consists of those pixels in V_n^q that have a larger value than a pre-specified threshold. This threshold can be set by the designer to leverage the concentration of the attraction region, in which a higher threshold indicates a more concentrated attraction region.

There are three basic operations in the typical convolutional neural network: (i) max pooling, which means that only the maxima of a small region is passed to the next layer; (ii) ReLU rectification, which is a nonlinear function $f(x) = max(x, \mathbf{0})$; and (iii) convolution, whose key parameters are its weight matrix W and bias vector b. In a deconvolutional neural network, the corresponding inverse operations are: (i) Unpooling: Though the max pooling operation is non-invertible, we can approximately invert it by recording the locations of the maxima in a set of switch variables when the input image is processed in the convolutional neural network. The value from the layer above is placed into the locations of the maxima according to the corresponding switch variable, such that the structure of maxima is preserved. These maxima are analogous to the salient information in the forward design communication framework, as salient information in the design is likely to be conveyed to human's perceptual processing units. (ii) Rectification: The approximate inverse operation of ReLU rectification is itself. (iii) Convolution: To approximately invert the convolution operation, the convolution operator with transposed weight matrix W^T is used.

4. Experiment

We conducted an experiment composed of four parts: (i) Estimate the feature representation of 110 vehicle images through a convolutional neural network, which shares the same structure as AlexNet and is trained by both the ImageNet and Flickr datasets. We use pretrained parameters obtained from a verified deep learning platform Caffe [Jia 2014]. (ii) Develop prediction models for the ten aesthetic attributes listed in Table 2. These attributes are used by design teams in the automotive industry [Burnap 2015]. Each prediction model is from the same feature representation to an aesthetic attribute. The value of an aesthetic attribute is obtained through the Markov Chain modeled in the crowdsourced human feedback as detailed in Section 4.1. (iii) Conduct feature selection based on the attribute prediction model. The L1 regresssion model allows us to select the influential predictors. Its regularization parameter α is adaptive through cross validation. (iv) Visualize the salient features selected from the previous step and empirically choose a threshold for visual attraction region that can reflect the desired concentration. In our study, we choose the threshold $\gamma = \mu + \sigma$, where μ is the mean pixel value in feature image and σ is the standard deviation of pixel values.

Table 2. Ten design attributes used for partial ranking evaluation for 2D vehicle images

Low Attribute	Awkward	Weak	Conservative	Basic	Conventional
High Attribute	Well Proportioned	Powerful	Sporty	Luxurious	Distinctive
Low Attribute	Passive	Traditional	Understated	Friendly	Mature
High Attribute	Active	Innovative	Expressive	Aggressive	Youthful
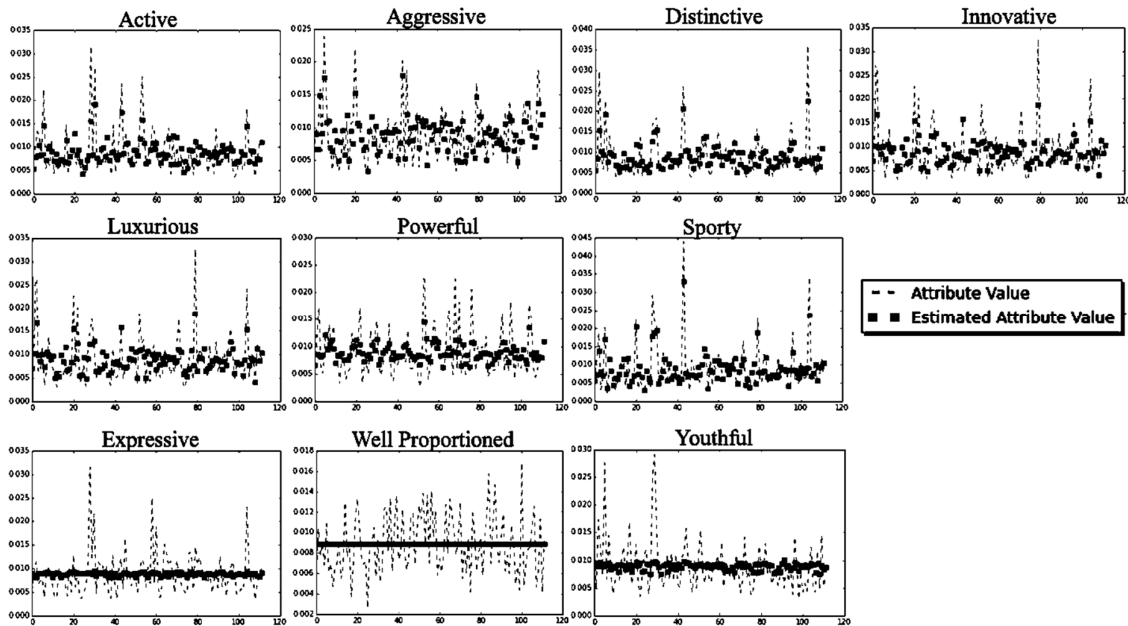
Figure 5. L1 regression prediction performance for all 10 design attributes with the x axis representing the vehicle ID and the y axis representing the attribute values and estimated values

4.1 Crowdsourcing for design attribute values

A databased-backed web application was developed to crowdsource partial rankings of the 110 vehicle images for the set of 10 design attributes from Table 2. These partial rankings were then aggregated using the Markov chain described in Section 3.2 to obtain the values of all 10 design attribute for each of the 110 vehicle models.

We gathered 361 participants through the crowdsourcing platform Amazon Mechanical Turk [Amazon Mechanical Turk 2014]. Participants were directed to an introduction page, where they were given instructions on ranking vehicles according to a semantic differential for a randomly assigned design attribute from Table 2. This semantic differential consisted of only one of the ten attributes from low to high value or vice versa to act as a counterbalance for ordering biases. Over the entire interactive survey, a participant was always given the design attribute semantic differential in the same direction (either "low value" to "high value" or vice versa) to reduce participant burden, though direction was randomized across participants. Next, participants were directed to the 2D design partial ranking page, with four vehicles chosen from the set of 110 vehicles in a top row and four outlined placeholders in a bottom row. Instructions on the page were given to drag-and-drop the four designs from the top row to the bottom row using the mouse, including the possibility of reordering the partial ranking.

5. Results and discussion

The attribute prediction performance is given in Figure 5. Seven out of ten prediction models provide attribute estimations that are similar to the attribute values obtained through crowdsourcing. This indicates that the features from the 5th layer in the AlexNet contain the important information for those seven attributes, and thus visualizing these features is a meaningful way to predict the attraction regions. Model fitness further validates the model for forward communication from the design to these seven attributes. However, the prediction model fails to predict three attributes including 'expressive', 'well-proportioned,' and 'youthful'. A possible reason is that the influence of the 5th layer features may not be as meaningful, and thus these three attributes are not included in our analysis and visualization.

Figure 6 shows the visual attraction region for the design attribute 'active'. We cover subsections of the design images with two groups of attraction regions. Each group corresponds to one salient feature. The images in the same row show the predicted attraction regions of the same feature for different cars. The predicted attraction regions focus on the same region of the car (front light in the first row) despite other

variations in the image space such as vehicle shape, color, and viewpoint. The images in the same column show the predicted attraction regions of different features in the same car. In this case, different attraction regions are shown for different features of the same car. It is important to point out that these attraction regions are estimated from our model without using eye-tracker data.
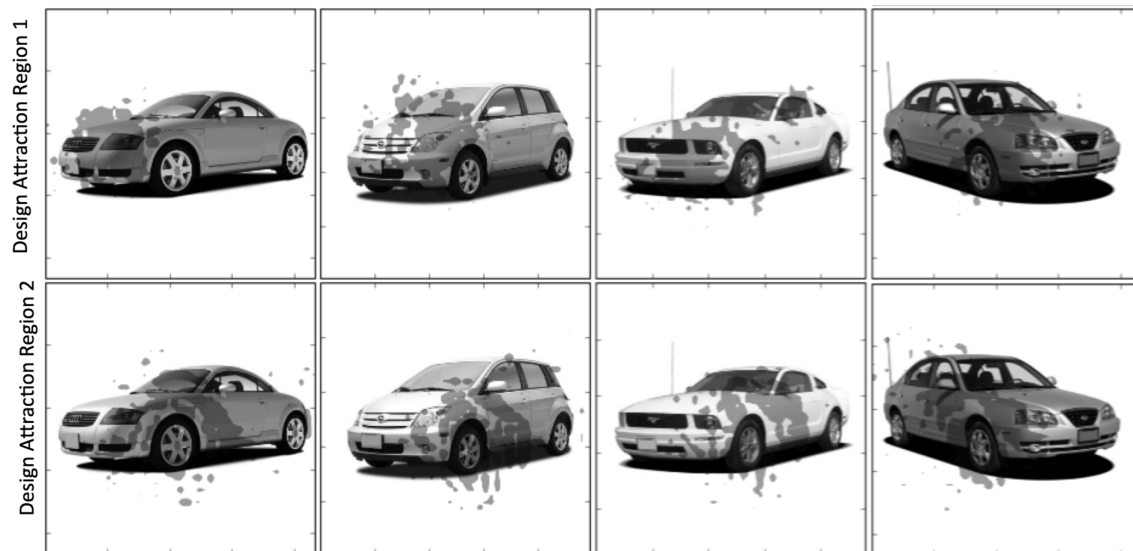
Figure 6. Examples of predicted attraction regions for design attribute 'Active'
The top row corresponds to an unknown design feature describing and `Active` car, seemingly focused on vehicle headlights, while the bottom row corresponds to a sepearte unknown design feature, seemingly focused on the front quarter-panel and door

5.1 Limitations and future work

A limitation of the present work is lack of validation. While qualitatively we can see that the predicted areas of visual attraction indeed only occupy subsections of the 2D vehicle images on the vehicle itself, we do not have an objective or independent measure. That is, while we capture a function in the forward direction from images to attribute values and visualize projections of its inverse from attribute features back to images, we do not have a way to assess whether the projected inverse mapping is correct. There are two difficulties here: (i) Defining an error metric for validity, and (ii) obtaining "ground truth" values for validity. Defining an error metric may be best addressed with assumptions from visual attraction models from psychobiological human attraction models.

While it may be possible to ask customers to give their "stated response" by clicking on regions of interest, it may be more fruitful to instead compare our predicted visual attraction with empirically-derived "revealed response" cues such as customer eye-tracking data [Du and MacDonald 2014], [Marshall et al. 2014], [Tovares et al. 2014] and implicit dimensionality reduction [Yumer et al. 2015]. One important future direction is better feature visualization to reveal more design details in the attention region. These design details may be valuable clues for designers to improve the aesthetic appeal of designed artifacts. Another future direction is the study of the approximately inverting process used in the model describing the backward communication, such as when it will fail and the bounds of errors. This is more theoretical work and beyond the scope of this paper.

6. Conclusion

Design visual attraction, in both descriptive and predictive senses, has a long history of study in the design community. The present work contributes a mathematical predictive model to identify visual attention regions that may attract the user's attraction.We introduced a data-driven method building on the framework of design as a communication process. We extended this method to include four stages in a forward - backward pipeline: (i) design feature learning, (ii) design attribute prediction, (iii) design

feature selection, and (iv) design feature visualization. This method is novel in that it is data-driven and does not require humans to provide the attention data. The modeling tools we used for the data-driven method include a convolutional neural network, L1 regression, a crowdsourced partial ranking Markov chain, and a deconvolutional neural network. This work is a first step toward data-driven predictions on how portions of the design space (regions of attraction) affect various design attributes via features learned using large-scale image data and selectively weighted crowdsourced perceptual responses. Future research to improve this method include validation using data-specific metrics, correlations with other methods such as eye-tracking that captures visual attraction in the design space, better feature visualization, and theoretical analysis of the proposed algorithm.

Acknowledgement

References

Amazon Mechanical Turk [website], Available at: <http://mturk.com/>, 2014, [Accessed 20.12.14].

Ben-Akiva, M., McFadden, D., Gärling, T., Gopinath, D., Walker, J., et al., "Extended framework for modeling choice behavior", Marketing Letters, Vol.10, 1999, pp. 187–203.

Berlyne, D. E., "Aesthetics and Psychobiology", Appleton-Century-Crofts, East Norwalk, CT, US, 1971.

Bloch, P. H., "Seeking the Ideal Form: Product Design and Consumer Response", Journal of Marketing, Vol.59, No.3, 1996, p. 16.

Brin, S., Page, L., "The anatomy of a large-scale hypertextual Web search engine", Computer networks and ISDN systems, Vol.30, 1998, pp. 107–117.

Burnap, A., Hartley, J., Pan, Y., Gonzalez, R., Papalambros, P. Y., "Balancing Design Freedom and Brand Recognition in the Evolution of Automotive Brand Character", In: Proceedings of the 2015 International Design Engineering Technical Conferences, Presented at the ASME Int. Design Eng. Technical Conf., Boston, MA, 2015.

Burnap, A., Ren, Y., Lee, H., Gonzalez, R., Papalambros, P. Y., "Improving Preference Prediction Accuracy with Feature Learning", In: Proceedings of the ASME 2014 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, New York, United States, August 17-20, 2014.

Buswell, G. T., "How people look at pictures: a study of the psychology and perception in art", 1935.

Chang, Y.-M., Chu, C.-P., Ma, M.-Y., "Exploring the Visual Cognitive Features on the Design of Car Based on the Theory of Eye-Tracking Technology", Przegląd Elektrotechniczny, Vol.89, 2013, pp. 143–146.

Coates, D., "Watches Tell More Than Time: Product Design, Information, and the Quest for Elegance", McGraw-Hill London, 2003.

Crilly, N., Moultrie, J., Clarkson, P. J., "Seeing things: consumer response to the visual domain in product design", Design Studies, Vol.25, 2004, pp. 547–577.

Dalal, N., Triggs, B., "Histograms of oriented gradients for human detection", In: CVPR 2005, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, pp. 886–893.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., Fei-Fei, L., "Imagenet: A large-scale hierarchical image database", In: CVPR 2009, IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248-255.

Du, P., MacDonald, E. F., "Eye-Tracking Data Predict Importance of Product Features and Saliency of Size Change", Journal of Mechanical Design, Vol.136, 2014.

Duchowski, A. T., "A breadth-first survey of eye-tracking applications", Behavior Research Methods, Instruments, & Computers, Vol.34, 2002, pp. 455–470.

Gatys, L. A., Ecker, A. S., Bethge, M., "A neural algorithm of artistic style", arXiv preprint arXiv:1508.06576, 2015.

Gerth, R. J., Burnap, A., Papalambros, P., "Crowdsourcing: A Primer and Its implications for Systems Engineering", Presented at the NDIA Ground Vehicle Systems Engineering and Technology Symposium, DTIC Document, Troy, MI, 2012.

Hubel, D. H., Wiesel, T. N., "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex", The Journal of Physiology, Vol.160, No.1, 1962, pp. 106–154.

Hubel, D. H., Wiesel, T. N., "Sensation and Perception", E. Bruce Goldstein University of Pittsburg and University of Arizona, 2014.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., et al., "Caffe: Convolutional architecture for fast feature embedding", In: Proceedings of the ACM International Conference on Multimedia, 2014, pp. 675-678.

Karayev, S., Trentacoste, M., Han, H., Agarwala, A., Darrell, T., Hertzmann, A., Winnemoeller, H., "Recognizing image style", arXiv preprint arXiv:1311.3715, 2013.

Kelly, J. C., Maheut, P., Petiot, J.-F., Papalambros, P. Y., "Incorporating user shape preference in engineering design optimisation", Journal of Engineering Design, Vol.22, 2011, pp. 627–650.

Krizhevsky, A., Sutskever, I., Hinton, G. E., "Imagenet classification with deep convolutional neural networks", In: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.

Lee, H., Ekanadham, C., Ng, A. Y., "Sparse deep belief net model for visual area V2", In: Advances in Neural Information Processing Systems, 2008, pp. 873–880.

Marčelja, S., "Mathematical description of the responses of simple cortical cells", JOSA 70, 1970, pp. 1297–1300.

Marshall, B. H., Sareen, S., Springer, J. A., Reid, T., "Eye tracking data understanding for product representation studies", In: Proc. of the 3rd Annual Conf. on Research in Information Technology, ACM, 2014, pp. 3–8.

McCormack, J. P., Cagan, J., Vogel, C. M., "Speaking the Buick language: capturing, understanding, and exploring brand identity with shape grammars", Design Studies, Vol.25, 2004, pp. 1–29.

Mono, R., "Design for Product Understanding: The Aesthetics of Design from a Semiotic Approach", Stockholm, Sweden: Liber AB, 1997.

Moulson, T., Sproles, G., "Styling Strategy", Business Horizons, Vol.43, No.5, 2000, pp. 45–52.

Nagamachi, M., "Kansei Engineering: a New Ergonomic Consumer-Oriented Technology for Product Development", International Journal of Industrial Ergonomics, Vol.15, No.1, 1995, pp. 3-11.

Norman, D. A., "Emotional Design: Why We Love (or Hate) Everyday Things", Basic books, New York, 2004.

Orsborn, S., Cagan, J., "Multiagent Shape Grammar Implementation: Automatically Generating Form Concepts According to a Preference Function", Journal of Mechanical Design, Vol.31, 2009.

Orsborn, S., Cagan, J., Boatwright, P., "Quantifying Aesthetic Form Preference in a Utility Function", Journal of Mechanical Design, Vol.131, 2009.

Orsborn, S., Cagan, J., Pawlicki, R., Smith, R. C., "Creating cross-over vehicles: Defining and combining vehicle classes using shape grammars", Artificial Intelligence for Engineering Design, Analysis, and Manufacturing, Vol.20, 2006, pp. 217–246.

Panchal, J. H., "Using Crowds in Engineering Design–Towards a Holistic Framework", In: Proceedings of the International Conference on Engineering Design, ICED 2015, Milano, Italy, 2015.

Petiot, J.-F., Dagher, A., "Preference-Oriented Form Design: Application to Cars' Headlights", International Journal on Interactive Design and Manufacturing, Vol.5, 2010, pp. 17–27.

Poirson, E., Petiot, J.-F., Boivin, L., Blumenthal, D., "Eliciting user perceptions using assessment tests based on an interactive genetic algorithm", Journal of Mechanical Design, Vol.135, 2013.

Pugliese, M. J., Cagan, J., "Capturing a rebel: modeling the Harley-Davidson brand through a motorcycle shape grammar", Research in Engineering Design, Vol.13, 2002, pp. 139–156.

Reid, T. N., Gonzalez, R. D., Papalambros, P. Y., "Quantification of Perceived Environmental Friendliness for Vehicle Silhouette Design", Journal of Mechanical Design, Vol.132, 2010.

Reid, T. N., MacDonald, E. F., Du, P., "Impact of product design representation on customer judgment", Journal of Mechanical Design, Vol.135, 2013.

Ren, Y., Burnap, A., Papalambros, P., "Quantification of Perceptual Design Attributes Using a Crowd", In: Proceedings of the 19th International Conference on Engineering Design. Seoul, Korea, 2013.

Ruckpaul, A., Fürstenhöfer, T., Matthiesen, S., "Combination of Eye Tracking and Think-Aloud Methods in Engineering Design Research", In: Gero, J. S., Hanna, S. (Eds.), Design Computing and Cognition '14. Springer International Publishing, Cham, 2015, pp. 81–97.

Tovares, N., Boatwright, P., Cagan, J., "Experiential Conjoint Analysis: An Experience-Based Method for Eliciting, Capturing, and Modeling Consumer Preference", Journal of Mechanical Design, Vol.136, 2014.

Yumer, M. E., Chaudhuri, S., Hodgins, J. K., Kara, L. B., "Semantic Shape Editing Using Deformation Handles", ACM Transactions on Graphics (Proceedings of SIGGRAPH 2015), Vol.34, 2015.

Zeiler, M. D., Fergus, R., "Visualizing and understanding convolutional networks", In: Computer Vision–ECCV, Springer International Publishing, 2014, pp. 818-833.

Yanxin Pan, PhD student
University of Michigan, Integrative Systems + Design
1301 Beal Ave, 48109 Ann Arbor, United States
Email: yanxinp@umich.edu