Eye Tracking for Screening Design Parameters in Adjective-based Design of Yacht Hull

Kemal Mert Dogan1*, Hiromasa Suzuki1 and Erkan Gunpinar2
1. The University of Tokyo
2. Istanbul Technical University

Abstract

Adjective-based design is a method that translates human perception into design parameters quantitatively in order to achieve better understanding between designers and clients. In this approach, adjectives are used to describe product designs, which are generated via design parameters in terms of geometry. As a requirement of the concept, relations between hull adjectives (e.g., comfortable and aesthetic) and design parameters (e.g., length and width) are learned via a machine-learning algorithm. Nevertheless, the relations cannot be represented by some of the design parameters, although they are in the learning process. This issue shows that the parameters do not impact the adjective choices but add noises to the learning process. Therefore, in this study, visual evaluations are made using eye tracking technology for screening the parameters based on their attractiveness and establishing relations between the attractive ones and the adjectives to enhance quality of the relation representations. Eye tracking is used in perceptual research, which proves the existence of correlations between gaze data and human preferences. The main advantage of eye tracking is that reliable human perception data can more likely be collected compared to the user tests, since the evaluation is based on subjects’ attention rather than applying solely questionnaires that are limited by the question content. In light of the benefits and finding, an eye tracking device is used to collect gaze data, and then, eye tracking tools such as Area of Interest (AOI), scan path, and heat map are used to evaluate attractiveness of the design parameters. Finally, regression analysis is used to represent relations between gaze data of design parameters and the adjectives.

Keywords: Industrial design, Human factor in design, Adjective-based design, Parametric design, Eye tracking, Generalized Linear Model (GLM)

1. Introduction

Adjective-based design

Product performance and qualities are not solely good enough criteria for clients as the market competition has already made them compulsory. For products with high quality and good performance, design is highly important for clients to make the final decision [1] [2]. However, this criterion is not a clear target to be reached and is measured through client satisfaction, which is highly related to human feelings; therefore, it is difficult to forecast this criterion beforehand. Dogan and Gunpinar [3] proposed the adjective-based design method, which quantizes human feelings and represents them as relations between adjectives (e.g., charismatic and modern) and design parameters that are

*Corresponding author. Email address: dogan@den.t.u-tokyo.ac.jp, Address: The University of Tokyo, Graduate School of Engineering, Bunkyo, Hongo, 7 Chome-3-1, 113-8654, Tokyo, JAPAN, Tel: +81-3-5841-6009.
geometrically defined on products (e.g., lengths and widths). Such approach aims to provide better communication between designers and clients [4]. In this method, feelings about designs can be easily expressed using adjectives and their relations with design parameters are used as a guide by designers. Thus, understanding clients and appealing them can be successfully achieved in relatively short times. Dogan and Gunpinar [3] applied the adjective-based method to yacht hulls. They first learned 10 hull adjectives using a survey method and defined design parameters on hull covering common hull features. A second survey was then applied to eliminate irrelevant parameters searching whether the modifications of these parameters affect any of the adjective choices made by the subjects. After the eliminations, the remaining design parameters were used to generate hull designs via a novel parametric design framework, which is also similar to the method proposed by Khan et al. [5]. Then, another survey was conducted, where the designed hulls were matched with the hull adjectives to create a dataset that included corresponding parameter values and the matched adjectives for each hull. Finally, a machine-learning algorithm that used neural network was fed with the dataset to represent relations between the adjectives and the design parameters as mathematical models. However, some design parameters’ relations with the adjectives could not be established; thus, it was concluded that the design parameters did not affect the subjects’ adjective choices. Such noises are caused by unrelated parameters and need to be removed to raise the quality of relation representations. In this study, it is assumed that if an appearance feature of a product does not attract subjects, the related design parameter with the feature is unattractive and irrelevant to the adjectives as well. An eye tracking device is used to evaluate the degree of attractiveness of design parameters, observing where the subjects’ attention is focused.

Eye tracking system

Eye tracking, where eye movements of subjects are recorded as gaze points on given stimuli, is a commonly used method in neuroscience, psychology, computer science, industrial engineering, and marketing [6]. In the eye tracking system, illuminators are used to generate near-infrared lights and images of their reflections from eyes are captured by cameras. In addition, an external processor is used to process the images to analyze the reflection patterns and generate a 3D model of the eye with its position in space. The 3D eye model is then used to calculate locations of the gaze points on relevant stimuli. Eye movements are illustrated on stimuli via fixations, which are a batch of gaze points accumulated within a predefined span in shape of circles, and saccades are displayed via straight lines that show eye movements from one fixation to another (Figure 1(a)). In this study, the attractiveness of target areas is analyzed through fixation locations and durations at the locations, as well as scan paths [7], which are combinations of fixations and saccades. Fixations are widely analyzed using the Area of Interest (AOI) tool and heat maps [8]. Using the AOI tool, targets are surrounded with diverse shapes (Figure 1(b)) to perform quantitative analysis with various metrics such as fixation duration and visit count. The heat map (Figure 1(c)) represents fixation-accumulated areas based on a color scale where red indicates more attractive areas than the rest of the image. One method to examine the scan path is defining AOIs (see Figure 1(d) for yellow rectangular shapes) on the path that needs to be searched and the transitions between the AOIs will indicate how often (frequency) the path is used. A transition matrix is also used to summarize the transitions in matrix form, as illustrated in Figure 1(e).

Template Learning Survey (TLS)

Design parameters can be classified as local and global parameters (Figure 2) according to their impact on hull appearance. Local parameters are the ones that have full control on the target shape such as radii of feature curves. On the other hand, global design parameters affect large areas (e.g., lengths and widths) and the recorded gaze data on these areas can be found in any shape related to the search behavior of subjects. Thus, gaze points for global parameters first need to be identified according to the search behavior of subjects in order to analyze them with an appropriate strategy. To do this, a survey is used, called TLS, where all the learned identifier information for the relevant parameter is named as template.

Proposed approach

According to the flow of the proposed approach displayed in Figure 3, the Eye Tracking Aided Survey (ETAS) is first applied, where participants select appropriate adjectives from a given list for the hull images and gaze points of participants are recorded. The local parameters are analyzed by defining AOIs, while the global parameters are first applied TLS to learn their templates to determine a suitable analysis method. If a template of a global parameter shows the gaze points in the shape of accumulation in a particular area, the parameter is also treated as the local parameter.
Dogan et al. (2018) 1–24

3

Figure 1: (a) Fixations are represented via circles and saccades are straight lines that connect consecutive fixations. (b) AOIs are defined via various shapes on target areas for further statistical analysis. (c) Heat map illustrates gaze point accumulations based on a color scale where red areas depict areas that receive maximum attention. (d) Scan path analysis via AOI transitions. (e) Transition matrix for scan path analysis.

Figure 2: Two cases to decide whether AOI will be defined on the relevant parameter or TLS will be applied to learn what the best strategy is.

since it is searched locally. On the other hand, if the gaze data are spread out instead of being gathered in local areas, scan path analysis is used. Relative time durations (RTDs) are then calculated for local parameters with the outcome of AOI analysis, and transition probabilities (frequency of path usage) are calculated for global parameters, which are analyzed via the scan path. In the end, the two calculated metrics for the design parameters are used to compare their attractiveness. Finally, relations between gaze data of attractive parameters and the responses (adjective choices) are searched using regression analysis.

2. Related Work

This study is related to the adjective-based design method, which is similar to other human-oriented design studies such as Kansei engineering [9] and semantic attribute-based concepts [10]. Since these design concepts put the human perception in the center of the design process, this study is also related to the cognitive research conducted to learn human preferences. We are specifically interested in the research where eye tracking systems are used.

Adjective-based design

Literature covers several studies on human preferences to explore, identify, and convert human emotions into physical entities to estimate their possible preferences [9]. For example, Yanagisawa et al. [11] created a model-based design method that used Kansei database to forecast customer preferences and obtain former understanding
Suggested that eye movements are related to the emotional states of subjects as they have diverse reactions for pleasant and unpleasant media. Yousefi et al. [14] suggested the use of eye tracking for construction processes, analyzing the relations between gaze data and user satisfaction. Mohammadpour et al. [15] applied user tests recording gaze data, where building designs were evaluated by subjects and long fixation durations for attractive designs were measured. Kim et al. [16] examined correlations between nightscape preference degrees and gaze points, and found that subjects spent more time on the preferred images. Noland et al. [17] used eye tracking in their survey, which allowed qualitative and quantitative investigations of images related to urban planning. The study results showed that cars and parking were ranked negatively, while new urbanist parts of the images such as people and pedestrian features were mostly highly ranked. This study is motivated to use eye tracking technology with these findings to evaluate the attractiveness of the design parameters.

**Eye tracking studies for product design**

There are a few studies that use the eye tracking framework to particularly understand outstanding design features or how human aesthetic understanding is formed by product appearance. Kukkonen [18] studied product designs with an example of mobile phones where gaze data of subjects were compared through a questionnaire and strong correlations were found between them. In that study, five mobile phones were shown in pairs to the subjects, including designers, and gaze data for product attitudes and preferences were analyzed. Köhler et al. [19] extended the well-known Kansei engineering method using eye tracking, in order to evaluate product aesthetics from the perspective of human perception, and validated the results using descriptive and statistical analysis procedures. Koivunen [20] claimed that the way people look at designs is important and suggested strategies about first impression of the subjects, which they thought can help designers. Du and MacDonald [21] also analyzed the relative importance of visual product attributes and the impact of the attributes’ sizes on human preferences by using an eye tracking device. Their study showed the existence of significant relations between the gaze data and attribute importance, and used ordinal logistic regression to predict the importance of both attributes and their size changes. Khalighy et al. [22] formulated aesthetic understanding of subjects through the gaze data depending on beauty and attractiveness drivers. They asserted that the aesthetic formula succeeds in estimating human preferences via gaze data.

The above studies aimed to learn important geometric features to help designers in understanding customer needs and desires. Our goal is in a similar direction; however, we go further and provide an adjective-based design concept which is a kind of translator that converts adjectives (which is the design aesthetic language of clients) into design parameters (which is the language of designers). Moreover, the importance of design parameters over appearance features is found, which are combination of design parameters. A design framework is also proposed that works with
the design parameters, which means that the findings can be directly used by the design framework unless subjective
decisions are required by designers. Besides all these benefits, our screening method is conducted more systematically
as an orthogonal array is used to sample designs having varied geometric features, where the design parameter values
are equally changed through the 54 hull designs. It is believed that to get favorable results from the design evaluations,
such systematic approach is needed to observe real perceptions of humans when they see varied features. To the best
of our knowledge, the existing product evaluation studies with eye tracking systems use limited number of existing
design samples, which may be good enough to propose their methods, but cannot offer proper evaluation due to lack
of systematically prepared enough design variations.

3. Experiment materials and setup

The experiments were conducted with 54 hull design images and 10 adjectives, which were obtained from Dogan
and Gunpinar [3]. In that study, frequently used adjectives were determined via a survey, where participants were
shown various yacht hull models and asked to describe them via adjectives. Social listening method was then used
to enlarge the adjective dictionary and the retrospective think aloud method [23] was used to validate the determined
adjectives via subjects’ feedbacks. Finally, the following 10 adjectives were determined: strong (A1), speedy (A2),
comfortable (A3), aesthetic (A4), usual (common) (A5), aggressive (A6), compact (A7), modern (A8), charismatic (A9) and
cute (A10). Before the surveys were conducted, each participant was provided car images related to the adjectives
for the illustration. The reason for selecting cars rather than yachts was to not steer the subjects while they were
matching adjectives with the hulls, but only show them how an adjective can be related to a product design. The car
images were obtained from the Web and are not shared with this paper due to copyright issues. In addition, dictionaries
were used to standardize adjective definitions for the subjects. In the end, each participant was trained with several
hull models for warming-up purposes.

Hull models were sampled using the Taguchi experimental method [24], which suggests an L54 orthogonal array
for 25 design parameters with three levels and one with two levels. Each level was assigned an appropriate value to
determine dimensions for the 54 hulls. The hull models were then generated via a novel parametric design framework
proposed by Khan et al. [5], where a hull is divided into three parts: entrance section, middle-body section, and run
section (Figure 4). According to the framework, the entrance section surface is obtained via loft operation between
forward profile (FP) and entrance station profile (SPe); middle-body section is generated by lofting between SPe and
middle-body station profile (SPm); and run section is generated by lofting between SPm and run station profile (SPr).
Note that the lofting processes between profiles were performed following top, bottom, and feature (middle) guide
curves to control the surfaces. Furthermore, only half of the model was first lofted and mirroring operation was then
performed to obtain the full model. The design parameters with their definitions are given in Table 1, and the hull
locations are illustrated in Figure 4.

Table 1: Design parameters with definitions.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>L²</td>
<td>Entrance section length</td>
<td>α</td>
<td>Deck line angle</td>
</tr>
<tr>
<td>Lᵐ</td>
<td>Middle-body section length</td>
<td>α₂</td>
<td>Backward angle</td>
</tr>
<tr>
<td>Lᵣ</td>
<td>Run section length</td>
<td>β</td>
<td>Entrance angle</td>
</tr>
<tr>
<td>Bᵣ</td>
<td>Entrance section width</td>
<td>θ</td>
<td>Forward profile angle</td>
</tr>
<tr>
<td>Bᵐ</td>
<td>Middle-body section width</td>
<td>H</td>
<td>Height of forward profile</td>
</tr>
<tr>
<td>Bᵣ</td>
<td>Run section width</td>
<td>Rₑ¹</td>
<td>Min. radius of entrance top guide curve</td>
</tr>
<tr>
<td>Dₑ₁</td>
<td>Entrance section depth</td>
<td>Rₑ¹</td>
<td>Min. radius of middle-body top guide curve</td>
</tr>
<tr>
<td>Dᵐ₁</td>
<td>Middle-body section depth</td>
<td>Rᵣ¹</td>
<td>Min. radius of run top guide curve</td>
</tr>
<tr>
<td>Dᵣ¹</td>
<td>Run section depth</td>
<td>Rₒ₁</td>
<td>Min. radius of entrance bottom guide curve</td>
</tr>
<tr>
<td>Dₑ₂</td>
<td>Entrance section upper depth</td>
<td>Rₑ²</td>
<td>Min. radius of middle-body bottom guide curve</td>
</tr>
<tr>
<td>Dᵐ₂</td>
<td>Middle-body section upper depth</td>
<td>Rᵣ²</td>
<td>Min. radius of run bottom guide curve</td>
</tr>
<tr>
<td>Dᵣ₂</td>
<td>Run section upper depth</td>
<td>R₅</td>
<td>Min. radius of upper profile curve</td>
</tr>
<tr>
<td>a₁</td>
<td>Width of bottom flat panel</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4: Design parameters are demonstrated on side, top, and station profiles, which are illustrated via A-A, B-B, and C-C cut views.

ETAS

ETAS questions were prepared using images of 54 hull models, with the front, top, and side views providing the same pixel/meter ratio based on lengths, widths, and depths to simplify the understanding of dimension changes and fair AOI definitions (Section 4.1.2). Note that additional views were not used since they would add nothing to the three views in terms of showing design parameters, but disperse the focus of the subjects to widen areas. Figure 5(a) shows the experimental setup, where the screen was 520 mm wide and 325 mm high (24 inch diagonal) and the image (one of the ETAS questions) had 1920 × 1080 resolution. A screen-based eye tracking device (Tobii Pro X3-120) was used with a sampling rate of 120 Hz. Moreover, participants were free to move their head within 500 mm width and 400 mm height according to the report of the producer. Owing to this setup, the subjects were not disturbed by the device, and hence, could behave more naturally. In addition, the tracker had a reported accuracy of 0.4° and precision of 0.24°. Finally, the distance between the participants and the tracker was set as 650 mm.

The ETAS question interface for a hull design is given in Figure 5(b), which shows the design with three views and the location of adjectives. Participants were asked to select at least two appropriate adjectives from the list in order to increase the time spent on the hull, as time limitation was not applied for ETAS. The responses were verbally expressed by the subjects and recorded by an operator via Google Form so that the subjects could keep looking at the screen and not interrupt the gaze recording. This part is important for computing the RTDs since each time a subject looks away from the screen to select an adjective leads to empty gaze data, while the time period of the questionnaire is increasing.

Before each experiment, calibration was made using a calibration tool of the tracker software using nine dots. ETAS was conducted in three sets, where each set consists of 18 hull design images and three calibration images (Figure 5(c)) between hull images with equal intervals. Note that the calibration images were shown only for two seconds and were used to check if the calibration was missed during the survey. Data of two subjects’ tasks were discarded from the analysis due to the calibration problems, which were detected via the calibration images.

Total number of participants for this study were 24, including two females, aged 20 and 42 with mean 24.21 and standard deviation 5.83. The nationalities of the participants were Chinese, Japanese, Korean, and Malaysian. Each participant was given only two tasks out of four (three ETAS sets and one TLS), and on average, 13 subjects.
participated in each ETAS set and 10 in TLS tasks. A relevant council of The University of Tokyo approved the surveys used in the study.

Figure 6 shows some gaze plots and heat maps of ETAS results obtained via Tobii Pro software [25].

Figure 5: (a) Experimental setup. (b) ETAS interface consists of front, side, and top views of a hull design and adjectives placed at bottom of the screen. (c) Calibration image.

Figure 6: (a) Gaze plots of some hulls. (b) Heat maps of some hulls.

4. AOI-based screening methods for design parameters

As mentioned in the first section, the local parameters are evaluated via the AOI tool and global parameters using scan path analysis. The scan path analysis was also conducted by defining AOIs on the path and the transitions between the AOIs refer the usage frequency of the defined paths. To prevent confusion, the AOIs used for scan path are called as path AOIs. In this section, first, design parameters will be classified as local and global based on their impact on the appearance and search areas using TLS. Then, AOI definitions and metrics that will be used will be explained.
4.1. Design parameter classification

Design parameters that have full control on the target shapes are classified as local parameters since their search area can be limited to particular areas without overlapping with other design parameters. These parameters, which have local impacts on target shape, include minimum radii of feature curves ($R_{1e}$, $R_{1m}$, $R_{0e}$, $R_{0m}$), bottom width of flat panel ($\alpha_1$), entrance angle ($\beta$), and upper depth parameters ($D_{2e}$, $D_{2m}$, $D_{2r}$). Note that the upper depth parameters are used to control the incline of the feature guide curve between the sections; therefore, the incline is analyzed via AOI to evaluate these parameters. In addition to these, FP parameters, which are $H$ and $\theta$, are also classified as local parameters to be evaluated via the AOI tool since the area covered by the profile is relatively small to use the scan path analysis. $H$ can be observed independently from all other parameters over the front view; however, over the side view, overlapping between $H$ and $\theta$ is inevitable due to sharing of the same geometric feature. Therefore, an AOI is defined with FP label that covers the whole FP on the side view; in other words, feature-based screening is performed for FP, which is a combination of $H$ and $\theta$, rather than parameter-based.

On the other hand, global design parameters include lengths ($L'$, $L^m$, $L^r$), widths ($B'$, $B^m$, $B^r$), and depths ($D_{1e}'$, $D_{1m}'$, $D_{1r}'$), since they impact large areas on the hull shape, and angular parameters ($\alpha, \alpha_2$), which can be found anywhere through the inclined lines. In brief, we do not know how global parameters are searched by subjects. Therefore, these design parameters first need to be identified based on their search area to determine the best strategy for them. TLS is suggested for the identification and to define templates according to the identities for each global design parameter. Here, the template means the generalized output of TLS for a design parameter that includes information of search area of the parameter, whether local or global, with size and location of the area. Note that all locally searched parameters are analyzed via the AOI tool even if they are classified as global design parameters due to their effects on the hull appearance.

4.1.1. TLS

TLS tasks are specialized to request participants search only one specific design parameter in order to record gaze data that belong only to this specific parameter. TLS tasks can be categorized into two types. The first task aims to compare a design parameter by sections of one hull model to observe the search behavior of subjects through the hull. Figure 7 demonstrates a question example of such task where relevant design parameters are first introduced and the task is explained without time limitation. Whenever participants feel ready, the calibration image is displayed for two seconds and a hull model is then shown for five seconds. Finally, a question (e.g., “Which section is shorter/longer than the others?”) is asked related to the task, and after the subject answers, the other hull model is shown. After that, the same task is repeated three more times with other hulls. This task is suitable for lengths, widths, and depths as they are defined for each section separately.

The second task is to compare a design parameter on three different hulls and suitable to be used for all design parameters. Figure 8 shows that the task is similarly explained and a calibration image again appears for two seconds, but this time, all hull models are shown consequently and one question (e.g., “Which model has wide/narrow/perpendicular angle?”) is asked to compare three hull models for the relevant parameter. The calibration images have extra goal in TLS, which forces participants to start at the same point to increase similarity of initial gaze points in order to get meaningful results for comparison. In addition, the question and answer options that are easy to remember and understand are chosen to enable participants easily focus on the main objective of the task. For example, sections are introduced as left, middle, and right for easy comparison rather than entrance, middle-body, and run, which are difficult to remember. Note that the answers given for the question tasks are not analyzed since their only aim is to give purpose to the subjects and do not add value for gaze data analysis.

All TLS results are first visualized by heat maps to check whether the parameters are searched locally (fixations are accumulated in particular area) or in a spread-out way, and additional tools are then used to decide if they are suitable for analysis via AOI or scan path. Example TLS outputs for globally searched design parameters are visualized by gaze plot and transition arrows between path AOIs, which are shown in Figure 9(a) for the first task and in Figure 9(b) for the second task. On the other hand, example local parameter results are given in Figure 10(a) for the first task and in Figure 10(b) for the second task. By evaluating the visual outputs, $L'$, $L^m$, and $L^r$ are classified as global parameters since they are searched horizontally in wide areas. Moreover, $B'$, $B^m$, $B^r$, and $\alpha_2$ are classified as local parameters since they are searched on particular areas where AOI definitions are possible without overlapping. Meanwhile, $D_{1e}'$, $D_{1m}'$, $D_{1r}'$, and $\alpha$ parameters cannot be classified due to random gaze plots, which are not suitable for both AOI and scan path analyses. Therefore, they are kept out of the scope of this paper.
Figure 7: TLS task flow of first type, where a question is asked after each hull is displayed to compare the hull sections’ relevant design parameter.

Figure 8: TLS task flow of second type, where a question is asked after three hulls are displayed to compare design parameters.

**Preliminary search via heat map**

The heat map is a well-known method to visualize aggregated gaze points of all subjects at a time on a stimulus using a color scale, as illustrated in Figure 1(c). This visualization method gives a general idea about the areas that are attractive to the subjects and those that are completely neglected. As Bojko [26] suggested, instead of using the heat map for extensive analysis, it is recommended to use as a visualization tool for planning the study, mainly to learn where to define AOIs in order to gain time and increase efficiency. Another issue that needs to be considered is the use of relative durations of the total time spent by the subjects in case the time periods of the experiments differ from one subject to another. The final suggestion is to set the total kernel size, which is the diameter of the colored area around the fixation center, with a larger value than the default to spread the colored area to provide higher confidence interval in terms of increasing the search area.

In our study, the heat maps are generated via Tobii Software by applying I-VT filtration [27] with default settings of Tobii Pro software [25] according to relative durations. Furthermore, although the default total kernel size of Tobii Pro software is 100 px, which is determined according to the highest visual span of the human eye, this value is set to 210 px to increase the confidence level of the investigated area. Finally, the parameters that need quantitative examinations by AOI in the colored areas are determined for each image.

As mentioned in Section 3, mirroring operation is used to obtain the final hull models, which means the hull model is symmetrically generated. Due to this reason, a half of the front and side views has the same design parameters as the other half. However, Figure 11 shows three heat map results, where the searching area of front and side views is indicated by arrow directions. According to this figure and TLS results (Figure 10), only half of the views are
investigated by the participants. Therefore, to simplify the study, unnecessary AOI definitions are avoided based on heat map visualizations and define the AOIs for half of the views only.

4.1.2. AOIs for local design parameters

Nineteen design parameters $R_{e1}^c$, $R_{m1}^c$, $R_{r1}^c$, $R_{e0}^m$, $R_{m0}^m$, $R_{r0}^m$, $R_4$, $R_5$, $D_{22}^e$, $D_{22}^m$, $D_{22}^r$, $\alpha_2$, $\beta$, $\theta$ (FP), $H$, $a_1$, $B'$, $B^m$, and $B^r$ are analyzed quantitatively via AOIs, since they are locally searched by the subjects. OGAMA [28] is used to define elliptical AOIs for $B'$, $B^m$, and $B^r$, since they are observed in an area close to ellipsoidal shape (Figure 10) and tetragonal AOIs for the remaining 16 design parameters. The AOI definition for parametrically generated designs has a critical problem that the dimension of the relevant design parameter differs over the 54 hull models, so the AOIs need to be arranged for each hull model parametrically. Three main issues are considered for AOI definition:
• **Fair comparison:** AOI sizes need to be adjusted by the same proportion for each of the 54 hulls, since comparison is performed between the parameters.

• **Batch AOI definitions:** Nineteen local design parameters for each of the 54 hull models require 1026 AOI definitions, which is impractical to do manually, so automation is needed to define AOIs.

• **Avoid from overlapping:** When AOIs are adjusted for each hull, there may be some circumstances that cause overlapped AOIs. To avoid this problem, the AOIs are defined on the center of the target geometric feature rather than the whole; in other words, the gaze points are sampled from the center of the target feature. This may cause some gaze data loss (decreases sensitivity), but prevents the analysis of incorrect gaze data (increases selectivity). As the design parameters are compared with batch AOI definitions, it is believed that such an approach is valid when AOI definition is made providing the fairness and necessity to define AOIs automatically without overlapping.

The volume of a 3D product design changes with its length, width, and depth parameters (basic parameters). Therefore, the AOI sizes are varied with equations, making them dependent on the main construction parameters according to the view. In other words, size of an AOI defined for a hull parameter on side view is based on length and depth parameters, while that of a top view parameter depends on width and length parameters. First, template elliptical AOIs are defined as covering the target areas, which are learned from TLS empirically. On the other hand, the templates of the tetragonal AOIs are defined via the oriented bounding box method [29], enclosing the target areas. After that, since the AOIs are managed via four handle points \( P_0, P_1, P_2, P_3 \), the distances between the handle points of the template AOIs in \( x \) and \( y \) directions are divided by the relevant basic parameters to compute ratios. Finally, the ratios are used to determine appropriate positions of the handle points for the 54 hulls. In other words, defined templates are reversed based on their hull models’ main dimensions and the handle points are relocated according to the new hull dimensions by the amount of the ratios. Note that a fixed reference point \( P_x^r \) is selected for each design parameter to compute the ratios and to relocate the handle points of the relevant AOIs.

There are several ways to define AOIs based on the basic parameters, but the one, that is used, is illustrated in Figure 12, where \( x \) and \( y \) coordinates of \( P_0, P_1, P_2, P_3 \) for the elliptical AOI of \( B^e \) is parametrically changed in proportion to the empirically found ratios via TLS over width \( (B^e) \) and lengths \( (L^1 \text{ and } L^2) \) according to reference point \( P_x^r \). Note that the major axis of the elliptical AOI is computed in two parts based on \( L^1 \text{ and } L^2 \) in proportions of the ratios, and the minor axis is computed based on \( B^e \). Similar and suitable approaches are determined and applied to the other local parameters to parametrically modify and enable them to orient with the various dimensions of the 54 hulls.

![Figure 12: AOI location and reference point \( (P_x^r) \) demonstration for \( B^e \).](image)

**4.1.3. AOIs for global design parameters**

As seen in Figure 9, the sequences of gaze fixations for the global parameters \( (L^e, L^m, \text{ and } L^l) \) are observed in the form of horizontal transitions along the hull on both side and top views. Our aim is to associate transitions made along
the sections with the relevant length parameter for each section. This can be done by defining path AOIs to construct a path similar to that observed in TLS and counting the transitions between the path AOIs. Therefore, eight path AOIs are defined: four on the side view, with labels $A_{1s}, A_{2s}, A_{3s},$ and $A_{4s}$, and four on top view, with labels $A_{1t}, A_{2t}, A_{3t},$ and $A_{4t}$ (Figure 13(a)). In addition to these eight path AOIs, two extra path AOIs are defined: one for the front view of the hull and one for the whole adjective region in order to be used for transition probability calculation (the reason for this will be explained later). Note that the distances $P_{0,1}^{s} - P_{2,3}^{s}$ of path AOIs are equal to the sum of maximum depth with cabin height on the side view and maximum width on the top view. Similar to the procedure for local parameters, the handle points of the AOIs in the $x$ direction are defined depending on length parameters and relocated through the 54 hull models using computed ratios. Furthermore, the OGAMA software provides average loci similarities in percentages, which depicts the proportion of the covered scan paths of the subjects by the defined path AOIs. The ratios for path AOIs are computed taking the loci similarities into consideration.

4.2. Evaluation metrics

Two metrics are used in this study: RTD for AOI analysis and transition probability for scan path analysis.

RTD

Since there are 54 hulls with varied geometries, the AOI sizes are parametrically adjusted for each hull. However, it is not fair to compare AOIs with different sizes [30]. Furthermore, the time periods for ETAS are not fixed, so the periods are different for each subject and the results can be dominated by the one who spends more time than the others. To address these problems, fixation durations are standardized to be relative to the time periods of subjects. Furthermore, the time periods for ETAS are not fixed, so the transition counts need to be normalized for comparison of the global parameters of the 54 hulls.

Equation 1 calculates the relative times ($\bar{T}^A$), where $T^c$ is the period of a subject on the relevant hull image and $T^A$ indicates the time spent on a defined AOI for a parameter.

$$\bar{T}^A = \frac{T^A}{T^c}$$ (1)

Equation 2 then normalizes the AOI size of the design parameter, where the sum of relative times of all subjects ($\sum_{subjects} \bar{T}^A$) is divided by AOI proportion of the design parameter to the whole question area ($k$). These computations are made for each of the 19 local design parameters to obtain the RTD ($S^A$).

$$S^A = \frac{1}{k} \sum_{subjects} \bar{T}^A$$ (2)

Transition probability

As in RTDs, the transition counts need to be normalized for comparison of the global parameters of the 54 hulls. To do this, the transition counts are divided by total transitions made in the relevant hull image and the results are interpreted as transition probability [31]. The general formula of transition probability ($p(A_i \mid A_j)$) for a pair of path AOIs ($A_i$ and $A_j$) is given in Equation 3, where the transition count ($C(A_i \parallel A_j)$) is divided by the total count of transitions made on the hull image ($C(A)$). Note that $A$ here refers to a relevant AOI label.

$$p(A_i \mid A_j) = \frac{C(A_i \parallel A_j)}{C(A)}$$ (3)

Since more than 26 target design parameters can be observed by subjects on given stimuli, many pointless transitions will be made for local parameters, which should be avoided from total transitions. Therefore, the path is constructed via 10 path AOIs covering the major transitions rather than the minors that are made on the image. $C(A)$ is the number of major transitions, which are the ones among front view, adjectives region, entrance section, middle-body section, and run section (sections are defined for both top and side views). Minor transitions are assumed as transitions made on black areas of the image and self-transitions ($A_i \parallel A_i$) are the ones made inside of the defined path AOIs, which are usually short saccades and more likely belong to the local parameters. Figure 13(a) shows the defined 10 path AOIs and Figure 13(b) displays the transition counts between the corresponding AOIs for both side and top views in the matrix form. Figure 13(c) illustrates the transition probability computation (Equation 3) over the transition matrix of the top view, where the transition probability between $A_{1t}$ and $A_{2t}$ ($p(A_{1t} \mid A_{2t})$) refers to that of the entrance section length ($p(L^e)$) on the top view, ($p(A_{1t} \mid A_{1t})$) is equal to $p(L^p)$ and $p(A_{1t} \mid A_{2t})$ is equal to $p(L')$. The
transition diagram in the Figure 13(c) is also used to summarize the probabilities between the relevant AOIs, where nowhere refers to AOIs other than \( (A_t^1, A_t^2, A_t^3, A_t^4) \). Similarly, transition probabilities of four path AOIs \( (A_s^1, A_s^2, A_s^3, A_s^4) \) are used to compute transition probabilities of the section lengths based on the side view. Front view and adjectives are other regions that are involved by the path, as they are important stations for the subjects.

Figure 13: (a) Ten path AOIs that cover major transitions. (b) A transition matrix example, which summarizes the transitions in matrix form. (c) Transition probability calculation example over top view path AOIs \{B, C, D, E\}, where the transition matrix members (transition counts) are divided by total transition count \( C(A) \). The probabilities between the path AOIs and remaining AOIs (nowhere according to investigated path AOIs) are then represented through a transition diagram.

5. Results

5.1. Screening-based Evaluation of Design Parameters

5.1.1. Evaluation with RTD metric

First, the ETAS questions were roughly examined based on five significant areas: front view, adjectives, entrance section, middle-body section and run section (the sections are available for both side and top views (Figure 14(a))), via their total RTDs computed for the 54 hulls \( \hat{S} = \sum_{i=1}^{54} S_{A} \). In Figure 14(b), \( \hat{S} \) values for each area are represented via a pie chart in percentile, which are relative to the sum of all. Note that percentages of sections were calculated according to sum of their computed durations on the side and top views. According to the chart, the regions ordered from the most attractive to the least attractive are as follows: 37% of the total RTD of all 54 hull images for the middle-body section, 20% for the entrance section, 19% for adjectives, 18% for front view, and 6% for the run section. This analysis suggests that the middle-body section is the first section that needs to be modified to appeal clients or achieve any design goals in terms of aesthetics, since people quickly detect the modifications made in this section. Also, the results show that dealing with run section modifications should be considered as last option since the clients will not pay attention to this section in the first place.

Figure 14(c) also shows the comparison of the view-based results with the adjectives. According to the comparison, the side view of the hull design attracts most attention, followed by the top view and then the front view, which means that the side silhouette of the designs need to be studied on priority to appeal clients.

Fine examination is performed with the 19 local design parameters via computed average RTDs \( \bar{S}_{A} \) over the 54 hulls. The distribution of \( \bar{S}_{A} \) values of the 19 parameters is represented via a box plot [32] in Figure 15(a). The box plot is composed of quartiles, which are used to evaluate the attractiveness of the parameters in four levels. Accordingly, the design parameters whose \( \bar{S}_{A} \) values are above than the third quartile (\( Q_3 \)) are the most attractive ones.
5.1.2 Evaluation with transition probability

 covered by them as compared to the rest. Figure 15(b), where the discarded parameters are surrounded by a red rectangle, indicating the smallness of the area.

 The results show that 25% of transitions are made for transition percentage for this parameter is 6 \( \bar{h} \) via scan path analysis and refer to the usage frequencies (or densities) of the relevant paths. The average different colors, where green demonstrates the most attractive parameters (Level 1), which are listed as follows:

 - **Level 1**: \( H, B', D_{2}^{m}, R_{4}, \) and \( R_{1}^{m} \).
 - **Level 2**: \( D_{2}', a_{1}, B^{m}, R_{5}, R_{0}^{m}, \) and \( FP \).
 - **Level 3**: \( \beta, R_{6}, R_{1}', \) and \( D_{2}' \).
 - **Level 4**: \( R_{1}', B', R_{6}, \) and \( a_{2} \).

 Since the attractiveness of the parameters is measured based on the gaze data, the relations between the gaze data and the hull adjectives must be analyzed. The idea is that if the gaze data have relations with the adjective choices, the attractiveness results obtained will be related to the adjective-based design concept, since it shows that the participants looked at the parameter to select the adjectives. Otherwise, the results will be unrelated to the adjective-based design concept. Therefore, Level 1 and Level 2 parameters are further analyzed via regression analysis to validate the attractiveness results. The parameters lower than the median value are (Level 3 and Level 4 parameters) kept out of the regression analysis. \( S^{A} \) proportions of each parameter relative to the sum of all \( (S^{A}/\sum_{i=1}^{19} S^{A}) \) are visualized in Figure 15(b), where the discarded parameters are surrounded by a red rectangle, indicating the smallness of the area covered by them as compared to the rest.

5.1.2 Evaluation with transition probability

 Transition probabilities of \( L', L_{m}' \), and \( L_{r}' \) parameters (\( p(L') \), \( p(L_{m}') \), and \( p(L_{r}') \)) are calculated for each of the 54 hulls via scan path analysis and refer to the usage frequencies (or densities) of the relevant paths. The average probabilities (\( \bar{p}(L'), \bar{p}(L_{m}'), \) and \( \bar{p}(L_{r}') \)) for the side and top views with sum of both are given by the bar chart in Figure 16(a). The results show that 2.5% of transitions are made for \( L_{m}' \) on the side view and 3.8% on the top view, so the total transition percentage for this parameter is 6.3% (\( \bar{p}(L_{m}') = 0.063 \)). The transition frequencies for \( L_{r}' \) on the side and top views are 1.6% and 1.3%, so the total value is 2.9% (\( \bar{p}(L_{r}') = 0.029 \)). Finally, \( L_{r}' \) transition frequency is calculated as 0.5% on the side view and 1.0% on the top view, and total frequency of the transitions compared to whole transitions made on the question area is 1.5% (\( \bar{p}(L_{r}') = 0.015 \)). It can be inferred from these results that the attractiveness order of lengths, in other words, priority of the length parameters for the adjective-based design concept is \( L_{m}' > L_{r}' > L' \).

 Figure 16(b) also depicts the distributions of the 54 hulls’ transition probabilities of \( L', L_{m}', \) and \( L_{r}' \) (sum of both side and top views for each hull model). In the box plots, each whisker shows the range where 25% of all probabilities are distributed, while the box covers the remaining 50%. Comparing to the range of box plots and median values of \( L', L_{m}', \) and \( L_{r}' \) (in order 0.03, 0.06, and 0.01), a big portion of \( p(L_{r}') \) values is distributed under \( p(L') \) and \( p(L_{m}') \) values. In addition, the run section has already been found unattractive; bringing all the reasons together, \( L' \) is also kept out from the regression analysis.
5.2. Relation Analysis via Regression Analysis

In this section, the relations between gaze data and adjective choices are studied. If some relations can be established, they can be used to find unattractive design parameters. For this purpose, regression analysis is applied for estimating the relations between the gaze metrics as independent input variables and adjective selections as dependent response variables. More specifically, the input variables are the RTD ($R^A$) and transition probability ($p$). They are denoted as $X = \{x_i, i = 1...13\}$. The response variables are 10 adjective selections {strong ($A_1$), speedy ($A_2$), comfortable ($A_3$), aesthetic ($A_4$), usual (common) ($A_5$), aggressive ($A_6$), compact ($A_7$), modern ($A_8$), charismatic ($A_9$) and cute ($A_{10}$)}. They are denoted as $Y = \{y_i, i = 1...10\}$. As shown in Table 2, for each of the 54 hull designs, relevant values for each of $x_i$ and $y_i$ are obtained with the user experiments described in Section 4. The regression analysis is conducted on these input and response variables with values from the 54 observations for the hull design ID = 1...54.

First, a dataset is prepared where input variables are assigned as gaze metrics and the response variables are defined by converting the adjective selections into levels. Next, Generalized Linear Model (GLM) [33] is introduced as the regression analysis method.

A set of input variables, called the regression model, is used for the regression analysis. The set X, including all 13 input variables, cannot be the best choice for the regression model but some subset of X is preferable. Thus, the best-subset selection method [34] is used to determine the final regression model for each adjective. In this method, the residual sum of squares (RSS) of all possible regression models is first calculated to determine the top models, which are then evaluated according to the Akaike Information Criteria (AIC) [35]. The AIC values of the top models are compared according to the Akaike weight ratios [36] and $D^2$ [37] values. Finally, the performance of the determined models for each adjective is measured via $D^2$, Mean Absolute Percentage Error (MAPE), and correlation analysis results.

Dataset organization and response variable determination

The gaze data are input for regression analysis and represented by RTD for 11 attractive local parameters and transition probability for two attractive global parameters. The adjective choices are converted into three levels according to the number of selections made by subjects, and then used as response variables for the analysis. Table 2 displays dataset organization including input variables, response variables, and observations.

Note that the levels represent the projected values of number of selections and will be analyzed as continuous variables rather than classes. The values of the gaze metrics, RTDs, and transition probabilities increase due to the
attention of the subjects. Accordingly, it is hypothesized that the number of selections of an adjective for a hull model is related to the increase in the metric values. In other words, if a design parameter is related to an adjective choice, the value of the relevant metric for the design parameter is increased by the number of participants who select the adjective. However, the amount of increase in the metrics and the number of subjects is not expected to be detected clearly, because each additional selection does not guarantee the increase in gaze metrics due to the unstable behavior of subjects during the survey. To clarify the increase between them, the number of selections are projected into levels, which provide higher jumps between the selections. Therefore, the increase will be easily related to the gaze data and the noises caused by ineffective selections are filtered.

Table 2: Dataset for regression analysis.

<table>
<thead>
<tr>
<th>Hull ID</th>
<th>Relative Time Durations</th>
<th>Transition Probabilities</th>
<th>Adjective Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_{rt}$ $S_{mt}$ $S_{rt}$ $S_{mt}$ $S_{rt}$ $S_{mt}$ $S_{rt}$ $S_{mt}$ $S_{rt}$ $S_{mt}$ $p(L)$ $p(L')$</td>
<td>$A_1$ $A_2$ $A_3$ $A_4$ $A_5$ $A_6$ $A_7$ $A_8$ $A_9$ $A_{10}$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.12 0.38 0.07 0.39 0.03 0.08 0.11 1.04 0.05 0.05 0.23</td>
<td>0.02 0.04</td>
<td>1 2 1 1 3 1 2 1 1 1</td>
</tr>
<tr>
<td>2</td>
<td>0.18 0.68 0.10 0.10 0.18 0.06 0.11 1.34 0.32 0.09 0.55</td>
<td>0.02 0.05</td>
<td>1 2 1 1 2 1 2 1 1 1</td>
</tr>
<tr>
<td>3</td>
<td>0.25 0.70 0.23 0.50 0.00 0.00 0.38 1.17 0.16 0.17 0.43</td>
<td>0.04 0.05</td>
<td>1 1 1 1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>4</td>
<td>0.10 0.04 0.03 0.40 0.00 0.26 0.30 1.02 0.00 0.50 0.40</td>
<td>0.03 0.04</td>
<td>2 1 2 2 1 1 1 1 1 1</td>
</tr>
<tr>
<td>5</td>
<td>0.10 0.23 0.08 0.56 0.00 0.03 0.17 1.05 0.46 0.41 0.12</td>
<td>0.02 0.05</td>
<td>2 1 1 1 1 1 2 1 1</td>
</tr>
<tr>
<td>6</td>
<td>0.12 0.43 0.03 0.24 0.05 0.16 0.36 1.97 0.20 0.25 0.48</td>
<td>0.02 0.06</td>
<td>2 2 2 2 1 1 1 1 1 1</td>
</tr>
<tr>
<td>7</td>
<td>0.13 0.07 0.17 0.11 0.27 0.19 0.09 1.15 0.25 0.11 0.10</td>
<td>0.02 0.03</td>
<td>1 2 1 1 1 1 2 1 1</td>
</tr>
<tr>
<td>8</td>
<td>0.09 0.20 0.04 1.14 0.03 0.11 0.00 0.82 0.19 0.99 0.04</td>
<td>0.03 0.06</td>
<td>2 1 2 1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>9</td>
<td>0.29 0.10 0.14 0.00 0.03 0.14 0.00 1.31 0.30 0.00 0.16</td>
<td>0.01 0.03</td>
<td>3 1 2 2 1 2 1 1 1 1</td>
</tr>
</tbody>
</table>

As it is mentioned before, each participant was given only two tasks out of four and some data were discarded due to calibration problems. Thus, the number of participants, namely the maximum possible number of selections, are different for each ETAS set, so the assignment of levels for the adjective choices must be done accordingly. Therefore,
instead of using the number of selections directly, selection indexes are calculated dividing them by the number of subjects participating in the relevant set. The 540 index values (54 hulls and 10 adjectives) are plotted against the hull ID using the histogram type graph to examine their distributions (Figure 17). The histogram is first divided into two parts by a median value of 0.46, and it is seen that the upper portion is randomly distributed. The stable part is then divided into two equal intervals. As a consequence, the levels are assigned using the following conditions:

- Level 1: $\text{index} < 0.23$
- Level 2: $0.23 \leq \text{index} < 0.46$
- Level 3: $\text{index} \geq 0.46$

Figure 17: Histogram of selection index values versus hull ID and red threshold lines for levels.

**GLM**

In this study, GLM [33] is used for the regression analysis. In addition, as the response variables (adjective levels) are non-negative continuous variables and limited between 1 and 3, which makes the distribution skewed since the bounds cannot be exceeded, the Gamma family [38] is found suitable for the data distribution after comparing and testing with the other families (e.g., Gaussian, Poisson, and Quasi). The general formula of GLM is as follows:

$$g(E(y_i)) = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k$$

where $E(y_i)$ is estimation of the response value (adjective choices) corresponding to $A_1, A_2, ..., A_{10}$; $x_1, ..., x_k$ are independent input variables (gaze metrics) corresponding to $S_{Rw}, S_{Bw}, S_{Bm}, S_{R4}, S_{R5}, S_{Rs}, S_{Fp}, S_{H}, S_{a1}, S_{D2}, S_{Dm}, p(L'), p(L^m)$; $\beta_0$ is intercept, and $\beta_1, ..., \beta_k$ are the coefficients for the estimation. Inverse link function ($g(\mu) = \mu^{-1}$) [38] is also used to relate the mean ($\mu$) of the Gamma distribution and the linear estimation model.

**Best-subset selection**

In general, a regression deviation consists of two parts:

$$\text{regression deviation} = \text{explained variance} + \text{residuals}$$

where variances are explained by the input variables (gaze metrics) and the residuals are remaining random variances that cannot be explained by the existing variables of the regression model. It is clear that our target is to increase the explained variances as much as possible, to decrease the residuals. To do this, the input variables, which are unrelated to the relevant response variable, need to be discarded from the regression model. This can be achieved via the subset selection method [34]. One way to select the best subsets is calculating the RSS via Equation 6 for each of the possible regression models, namely each $\hat{X}$ of the power set of $X, 2^X$. After that, the first $m$ best subsets, which
have the least RSS values, are found using an exhaustive search algorithm proposed by Morgan and Tatar [39]. RSS is computed for each $\hat{X}$ of $2^k$ as follows:

$$RSS = \sum_{i=0}^{n}(y_i - \hat{y}_i)^2$$

where $n$ is the number of observations (54, the number of hull designs), $y_i$ is the observed response value, and $\hat{y}_i$ represents the estimated value via the subset $\hat{X}$. $\{\hat{X}_1, \hat{X}_2, ..., \hat{X}_m\}$ can be found having the smallest RSS values. Then, $\{\hat{X}_1, \hat{X}_2, ..., \hat{X}_m\}$ are further analyzed according to one of the quality measures, i.e., AIC [35] value:

$$AIC = -2 \times \ln(L) + 2k$$

for each $\hat{X}_i \in \{\hat{X}_1, \hat{X}_2, ..., \hat{X}_m\}$. $L$ is the value of maximized likelihood function $L(\psi; x)$, which is calculated by setting its derivative (with respect to unknown parameter $\psi$) to zero and solving the derived equation for $\psi$; $k$ is the cardinality of $\hat{X}$. Since the AIC value is a measure of information loss, the model with the smallest AIC is the best. On the other hand, in case the AIC values are close to each other, another model close to the best model (having AIC value close to smallest) can be selected if it explains more portion of variance of the data. $D^2$ (Equation 8a) [37] is a measure of the explained variance for a GLM model, which is also equivalent to the well-known measure $R^2$. The value of $D^2$ increases with the number of variables involved in the regression model and causes unfair comparison between the models having different numbers of variables. Therefore, the regression models are compared via adjusted $D^2$ (Equation 8b) [37] to eliminate this problem.

$$D^2 = 1 - \frac{\text{Residual Deviance}}{\text{Null Deviance}}$$

$$\text{adjusted } D^2 = 1 - \frac{(n - 1)}{(n - k)} \times [1 - D^2]$$

where Residual Deviance refers to the residuals of the model including all model variables ($\beta_0 + \beta_1 x_1 + ... + \beta_k x_k$) and Null Deviance is the residual of the intercept-only model ($\beta_0$). Additionally, the AIC values of the models are compared via the following equation [36]:

$$w_{ij} = \frac{L_i}{L_j} \times (k_j - k_i)$$

where $w_{ij}$ is the ratio of Akaike weights [40], which indicates how much $i^{th}$ model $\hat{X}_i$ is better in terms of minimization of the information loss over $j^{th}$ model $\hat{X}_j$.

To sum up, the subset selection starts with finding several best-subsets, which have the least RSS values. AIC is then used to find the best among them, which has the smallest AIC value. The best subsets having AIC values close to the best model are then checked to see if their $D^2$ values are higher than those of the best model without dramatically raising $w_{ij}$, which is computed between the best model (which has minimum AIC value) and the inquired model. In the end, if the increase in $D^2$ value is not sufficient to afford the increase in $w_{ij}$, the search is concluded.

Performance results of determined regression models

To conduct the GLM analysis supporting the subset selection method based on RSS and AIC values, the best-subset GLM (bestglm) package is used proposed by McLeod and Xu [41] through R-Studio software [42]. Table 3 shows the summary of the results for the determined model for each adjective in matrix form, where the first column represents the adjective labels. The second column indicates MAPE, which is computed via Equation 10 to obtain the error between the estimated ($\hat{y}_i$) and observed ($y_i$) response values scaling from 0 to 1, where 0 refers to 100% prediction accuracy.

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$
Next, four columns represent correlation analysis results, which are measured with 95% confidence level, between the relevant adjective response of the data and fitted values of the models, where cor represents correlation value, $CI_L$ and $CI_U$ are lower and upper bounds of the confidence interval, and $p - value$ is significance level of the correlation. Furthermore, GLM assumes that all input variables are independent. This assumption is checked via a multicollinearity test where the generalized variance inflation factor (GVIF) [43] values are used for the judgment. The last row shows that the computed GVIF values for each variable are less than 3, which indicates low correlation between the input variables, namely their independency [44].

The results show that the ratio of explained variances ($D^2$), which can be interpreted in percentile multiplying them by 100, ranges between 16% and 45%. Since the human perception is a complicated issue, the collected data for social research are usually very noisy and the $D^2$ values even under 50% are acceptable in the context of the research [45]. Also note that, the regression analysis in this study is conducted to show the existence of correlation between gaze data and preferences, but not for the prediction purposes. Moreover, the computed MAPE values range between 0.16 and 0.39, which shows relatively good results to claim the existence of the relations considering the noisy data. Correlation values, where “−1” indicates perfect negative correlation and “0” depicts no correlation while “1” represents perfect positive correlation between the models and the responses, are found between 0.37 and 0.67, which refer to positive correlations between the gaze data and the preferences. The correlations for the $A_4$, $A_5$, and $A_{10}$ models are weak compared to the others, but still show the existence of positive relations. The $p - values$ also indicate that all correlations are significant since all of them are lower than the 0.05 value (according to 95% confidence level).

The overall performance results of the determined regression models validate the existence of relations between human preferences (adjective choices) and subject attention (gaze data). The parameters of the found regression models (determined $\hat{X}$ subsets for each adjective) are also represented in Table 3 with black color, which are highly related to the relevant adjectives in terms of attention. For example, the regression model variables of $A_1$ are $R_0^m$, $R_0$, $FP$, $H$, $D_2^c$, $D_0^m$ and $L^m$. A designer can use these significant parameters in the first stage to get the hull model according to the desired adjective by modifying an existing hull model as they are attractive and have impact on the adjective choices.

### Table 3: Regression analysis results and the related parameters with the relevant adjectives represented by black cells.

<table>
<thead>
<tr>
<th>Adjective</th>
<th>$D^2$</th>
<th>MAPE</th>
<th>cor</th>
<th>$CI_L$</th>
<th>$CI_U$</th>
<th>$p - value$</th>
<th>$R_0^m$</th>
<th>$R_0$</th>
<th>$R_0^m$</th>
<th>$FP$</th>
<th>$H$</th>
<th>$D_2^c$</th>
<th>$D_0^m$</th>
<th>$B^m$</th>
<th>$B^m$</th>
<th>$L^m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>0.38</td>
<td>0.31</td>
<td>0.62</td>
<td>0.43</td>
<td>0.76</td>
<td>5.02E-07</td>
<td>$R_0^m$</td>
<td>$R_0$</td>
<td>$R_0^m$</td>
<td>$FP$</td>
<td>$H$</td>
<td>$D_2^c$</td>
<td>$D_0^m$</td>
<td>$B^m$</td>
<td>$B^m$</td>
<td>$L^m$</td>
</tr>
<tr>
<td>$A_2$</td>
<td>0.25</td>
<td>0.39</td>
<td>0.51</td>
<td>0.28</td>
<td>0.69</td>
<td>7.54E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_3$</td>
<td>0.45</td>
<td>0.31</td>
<td>0.67</td>
<td>0.49</td>
<td>0.80</td>
<td>2.60E-08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_4$</td>
<td>0.16</td>
<td>0.23</td>
<td>0.37</td>
<td>0.11</td>
<td>0.58</td>
<td>6.96E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_5$</td>
<td>0.20</td>
<td>0.31</td>
<td>0.38</td>
<td>0.13</td>
<td>0.59</td>
<td>6.96E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_6$</td>
<td>0.35</td>
<td>0.31</td>
<td>0.60</td>
<td>0.39</td>
<td>0.75</td>
<td>1.73E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_7$</td>
<td>0.38</td>
<td>0.20</td>
<td>0.57</td>
<td>0.36</td>
<td>0.73</td>
<td>5.46E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_8$</td>
<td>0.30</td>
<td>0.32</td>
<td>0.54</td>
<td>0.31</td>
<td>0.70</td>
<td>2.83E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_9$</td>
<td>0.41</td>
<td>0.16</td>
<td>0.62</td>
<td>0.43</td>
<td>0.76</td>
<td>4.93E-07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_{10}$</td>
<td>0.22</td>
<td>0.28</td>
<td>0.42</td>
<td>0.17</td>
<td>0.62</td>
<td>0.001735</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note that the unattractive parameters determined in Sections 5.1.1 and 5.1.2 were also analyzed via plots against residuals of the regression models (Figure 18) in order to check if the parameter can explain some portion of the residual part of the model. However, the residual plots show that the parameters have still totally random distributions, which means they are not explanatory, or the randomness is broken locally but the performance measures are not raised to be involved in the regression models. For example, the plot for $R_0^m$ and $\beta$ is relatively symmetric according to the line $y = 0$, so it is not possible to estimate the residuals via them. On the other hand, in the plot of $R_0^c$, some residuals can be estimated by the increase in this parameter. However, after adding the parameter to the relevant model, the adjusted $D^2$ value decreases and MAPE value increases. Based on this, it can be said that the previous determination of the unattractive parameters is valid since they failed to explain the variances.
Further analysis

Pearson’s product moment correlation [46] is also used between time spent on each adjective (each adjective was surrounded by an AOI separately) and the adjective levels (Figure 17). Such analysis can provide following information:

- **How comfortable the subjects are with the adjectives:** If the subjects spend more time on an adjective even it is not selected, negative correlation is expected, which means that the adjective is difficult to match with the hulls or somehow confusing, which makes the subjects think more about it. So, in that case, it can be concluded that it needs to be replaced with a better adjective in the future surveys.

- **Consistency of gaze data and selection:** Reliability of the responses can be confirmed in cases where positive correlations are observed. However, negative correlations do not indicate bad reliability.

Table 4: Correlation results for adjectives based on durations

<table>
<thead>
<tr>
<th>Correlation Value</th>
<th>p &lt; .05</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>strong</td>
<td>0.37</td>
<td>0.0059</td>
</tr>
<tr>
<td>speedy</td>
<td>0.23</td>
<td>0.0905</td>
</tr>
<tr>
<td>comfortable</td>
<td>0.42</td>
<td>0.0014</td>
</tr>
<tr>
<td>aesthetic</td>
<td>0.32</td>
<td>0.0202</td>
</tr>
<tr>
<td>usual</td>
<td>0.41</td>
<td>0.0023</td>
</tr>
<tr>
<td>aggressive</td>
<td>0.44</td>
<td>0.0010</td>
</tr>
<tr>
<td>compact</td>
<td>0.43</td>
<td>0.0013</td>
</tr>
<tr>
<td>modern</td>
<td>0.25</td>
<td>0.0708</td>
</tr>
<tr>
<td>charismatic</td>
<td>0.47</td>
<td>0.0003</td>
</tr>
<tr>
<td>cute</td>
<td>0.08</td>
<td>0.5898</td>
</tr>
</tbody>
</table>

Table 4 shows the correlation results for the adjectives. According to the results, the correlations for *speedy*, *modern*, and *cute* are found insignificant as the *p* – *values* are lower than 0.05, which means the gaze data for them are random relative to the selections. Therefore, nothing can be claimed about them. On the other hand, remaining adjectives have weak positive correlations (0.32-0.47). Considering the noisy data that make the correlations weak, it still validates that the adjectives are perceived by the subjects and they find the adjectives suitable to match with the hulls.

Several correlation analyses are also conducted based on the time to first fixation (TTFF) metric, which indicate the time spent before the subjects’ first attention for a target AOI. The analyses are made for the adjective selections to inquire their relations with each adjective’s AOIs and with local design parameters. However, significant correlations are not found, which means that subjects do not have specific parameter targets related to the adjectives to
search in the first place, but they examine the designs until finding meaningful reasons for their choices. On the other hand, the relation between TTFF and RTD metrics of local design parameters is also checked and find interestingly significant correlations for \( R_4 \) (cor = 0.40; \( p \leq 0.05; t = 3.1890 \)), \( R^m_0 \) (cor = 0.33; \( p \leq 0.05; t = 2.5578 \)), \( FP \) (cor = 0.52; \( p \leq 0.05; t = 4.3646 \)) and \( a_1 \) (cor = 0.28; \( p \leq 0.05; t = 2.0723 \)). The existence of the correlation between RTD and TTFF for the parameters is also represented by graphs in Figure 19 with relation equations and \( R^2 \) values (represents how much variance is covered by the equation). Although it cannot be claimed that they are more significant than the other parameters based on these results, it is apparent that they require special attention in the design process.

### Table 5: Correlation results for adjectives based on durations

<table>
<thead>
<tr>
<th>Correlation Value</th>
<th>p &lt; .05</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R^m_1 )</td>
<td>-0.19</td>
<td>0.1678</td>
</tr>
<tr>
<td>( R_4 )</td>
<td>0.40</td>
<td>0.0024</td>
</tr>
<tr>
<td>( R_5 )</td>
<td>0.00</td>
<td>0.9864</td>
</tr>
<tr>
<td>( R^m_0 )</td>
<td>0.33</td>
<td>0.0135</td>
</tr>
<tr>
<td>( FP )</td>
<td>0.52</td>
<td>0.0001</td>
</tr>
<tr>
<td>( H )</td>
<td>-0.19</td>
<td>0.1757</td>
</tr>
<tr>
<td>( a_1 )</td>
<td>0.28</td>
<td>0.0432</td>
</tr>
<tr>
<td>( D^m_2 )</td>
<td>-0.10</td>
<td>0.4850</td>
</tr>
<tr>
<td>( B^m )</td>
<td>-0.01</td>
<td>0.9149</td>
</tr>
</tbody>
</table>

### Figure 19: TTFF versus RTD of \( R_4, R^m_0, FP \) and \( a_1 \)

6. **Conclusion and Discussion**

In this study, the adjective-based design concept is used, where hulls are parameterized and the adjectives are related to the designs at the parametric level. This provides relatively more detailed examination compared to the studies conducted with appearance features of designs, which are obtained by a combination of varied values of relevant design parameters. As the design elements are correlated to the targets at the parametric level, such approach
provides flexibility and more control on sampling designs that meet the targets, specifically after the outcomes of this perception study are integrated into the design concept. However, working with the design parameters requires more attention as different dimension values and various combinations of the design parameters can be interpreted differently by subjects. Therefore, the Taguchi experimental method is also used to sample models, where each design parameter value is shown to be able to cover various features equivalently in the same number of times.

TLS is first proposed, which is used to classify the global design parameters based on search behavior of the subjects. One output of TLS is a template, which conveys the required information about AOI definitions based on search areas of the relevant design parameters. An ETAS method was then utilized to collect gaze data, which were evaluated via RTD and transition probability metrics. The evaluations were conducted in three parts: rough examination (section based and view based), fine examination (design parameters), and finally, regression analysis. Rough examinations indicated the most attractive section as middle-body section and view as side view. Such results can be used by designers to decrease labor times working only on the attractive areas. Fine examinations were first used to determine attractive parameters in four levels for local parameters and priorities for the global parameters interpreting a box plot. The determined attractive parameters for fine examination results were then used in the regression analysis to validate that attractive parameters are really useful for the adjective-based design concept.

GLM with the best-subset selection method was used to find regression models for each adjective depending on some performance measures ($D^2$ to measure explained variance ratio and MAPE for accuracy check) and best-subset $\hat{X}$ determination criteria (RSS and AIC). In addition, correlation analyses were conducted to confirm and measure levels of relations. The results proved the existence of correlation between gaze data and adjective choices, even though the performance measures were relatively low compared to the predictive models. However, our aim was not to get predictive models but to show the relations. Social science is complicated to study, since the human understanding is difficult to estimate and this problem is getting even more complex with the unclear and varied targets. 54 yacht hulls are studied, which are generated by 26 design parameters, and developed 26 targets from our viewpoint; however, there are even more targets for the subjects as a hull design has many more design details that can be detected. Moreover, they are asked to match 10 adjectives with 54 hulls having all the design details, that makes the problem even harder, since the decision-making mechanism is also developed and changed by seeing various hull models and unstable decisions (noises) are observed besides the useful data. Moreover, they are not designers, who change their decision criteria for no tangible reasons. After all of this, despite the weakness of the performance measures compared to the predictive models, significant results are found that validate the attractiveness of the design parameters determined by the gaze metrics. Specifically, relevant design parameters, which are variables of determined subsets, have priorities to get desired adjective-based design among all attractive design parameters, which can be used by designers.

The results are suggested to be combined with the human-oriented design [3] and parametric design [5] approaches to generate an infinite number of design options related to the defined adjectives. After that, the product-related requirements can be considered for filtering out the undesired design options. In this study, a yacht hull was used, in which the hydrodynamic and hydrostatic performances are crucial. As such performance properties in addition to aesthetic properties are important, the aesthetics are handled first to be then combined with a performance-based system, which is considered as future work.

Besides all benefits of the proposed method, design parametrization and uniform design sampling are compulsory issues to benefit from the outcomes of this study. Otherwise, the design parameter screening cannot be accomplished thoroughly and misleading results are more likely to be obtained. We find it valuable to propose a technique that parametrizes designs and evaluates the parameters with automatically defined AOIs based on the parametrization. Doing that, we would like to encourage industrial design area to use our method to get prominent designs with even less labor, time, and costs.

Acknowledgment

We would like to thank Dr. Koichi Ohtomi for his technical discussions and Hiroyuki Katayama for his help on experimental setup. This paper is partially based on results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).
References

[24] G. Taguchi, G. Taguchi, System of experimental design; engineering methods to optimize quality and minimize costs, 1988.
[35] G. Taguchi, G. Taguchi, System of experimental design; engineering methods to optimize quality and minimize costs, 1988.
[46] K. Pearson, Mathematical contributions to the theory of evolution.---on a form of spurious correlation which may arise when indices are used in the measurement of organs, Proceedings of the royal society of london 60 (1896) 489–498.