Computational Aesthetic Evaluation of Logos

JIAJING ZHANG and JINHUI YU, Zhejiang University
KANG ZHANG, The University of Texas at Dallas
XIANJUN SAM ZHENG, Tsinghua University
JUNSONG ZHANG, Xiamen University

Computational aesthetics has become an active research field in recent years, but there have been few attempts in computational aesthetic evaluation of logos. In this article, we restrict our study on black-and-white logos, which are professionally designed for name-brand companies with similar properties, and apply perceptual models of standard design principles in computational aesthetic evaluation of logos. We define a group of metrics to evaluate some aspects in design principles such as balance, contrast, and harmony of logos. We also collect human ratings of balance, contrast, harmony, and aesthetics of 60 logos from 60 volunteers. Statistical linear regression models are trained on this database using a supervised machine-learning method. Experimental results show that our model-evaluated balance, contrast, and harmony have highly significant correlation of over 0.87 with human evaluations on the same dimensions. Finally, we regress human-evaluated aesthetics scores on model-evaluated balance, contrast, and harmony. The resulted regression model of aesthetics can predict human judgments on perceived aesthetics with a high correlation of 0.85. Our work provides a machine-learning-based reference framework for quantitative aesthetic evaluation of graphic design patterns and also the research of exploring the relationship between aesthetic perceptions of human and computational evaluation of design principles extracted from graphic designs.

1. INTRODUCTION

Graphic design is the art of communication, stylizing, and problem-solving through the use of graphics. Common uses of graphic design include identity, logos, publications, print advertisements, posters, billboards, website graphics and elements, signs, and product packaging. Graphic design contains at least two parts: the composing elements and the spatial arrangement of those elements. The second part is usually guided by design principles, in which balance, contrast, harmony, variety, rhythm, movement, proportion, and emphasis [White 2011] are commonly acknowledged.

In the past, aesthetic evaluation of graphic design was done by human experts qualitatively in words in an ad hoc fashion. Without quantitative evaluation tools, designers may take years to master design principles through painful design practice, and novices may face difficulties trying to tap into design aesthetics. Besides, it is practically impossible to ask experts to evaluate every new logo design or exhaustively compare similar designs.

Computational aesthetics is the research of computational methods that can make applicable aesthetic decisions in a similar fashion as humans can [Hoenig 2005; Fishwick 2008] and has become an active research field in recent years. The automated and quantitative evaluation capability of graphic designs would not only benefit to designers during their learning phase but also to the target audience or customers who have little knowledge of graphic design. It has a profound impact on the computer-aided study of design aesthetics, advanced search for design patterns guided by design principles, and identification of design styles.

A logo (abbreviation of logotype) is a graphic mark, emblem, or symbol commonly used by commercial enterprises, organizations, and even individuals to aid and promote instant public recognition. Logos are either purely graphic (symbols/icons) or are composed of the name of the organization (a logotype or wordmark), and here we show some examples of black-and-white graphic logos in Figure 1. The logos selected in our study given in Figure 8 include monochrome and color type; to rule out the influence of the color factor, we convert all these logos into black and white. This makes it easier to evaluate a few important aspects of logos, such as balance, contrast, and harmony.

Logo design is an important area of graphic design, and one of the most difficult to perfect. On one hand, design artists may use all design principles with varying degrees to organize individual elements into a workable, aesthetic design space in logos. On the other hand, common audiences know much less about design principles than design artists do and might be just aware of a few familiar aspects in logos, say, balance, contrast, and harmony.

Computational aesthetic evaluation of logos is challenging, and there have been few attempts in this area, and also there have been few attempts in applying standard design principles in computational aesthetic evaluation [Galanter 2012]. The challenges include the following aspects:

- Design principles itself is a big concept instead of simple visual features. Many factors influence each design principle. Previous works in evaluating design principles such as subject-background contrast in photographs [Wong and Low 2009] and color harmony in paintings [Li and Chen 2009] cannot be used in aesthetic evaluation of logos, since they do not consider the shape factor in aesthetic evaluation. Although we define each design principle in logos in a specific aspect, it is difficult to quantitatively measure such a design principle in terms of a human's judgment.
- Frameworks of computational aesthetic evaluation of web pages, photographs, and painting [Li and Chen 2009; Dhar et al. 2011; Obrador et al. 2012; Reinecke et al. 2013] are not entirely suitable for aesthetic evaluation of logos that follow different design principles.

In this study, we focus on developing quantitative models for measuring aesthetics of logos. We define a group of metrics to evaluate some aspects in design principles such as balance, contrast, and harmony
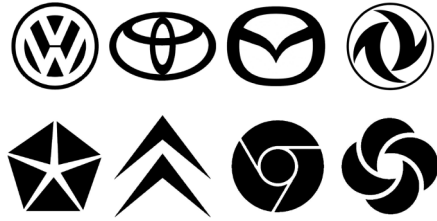
Fig. 1. Examples of black-and-white logos.

of logos. We also conduct a survey to collect human ratings of balance, contrast, harmony, and aesthetics of 60 logos from 60 volunteers. Utilizing these data, we develop regression models of balance, contrast, and harmony that have a highly significant correlation of over 0.87 with human evaluations on the same dimensions. Finally, we regress human-evaluated aesthetics scores on model-evaluated balance, contrast, and harmony. The resulted regression model of aesthetics can predict human judgments on perceived aesthetics with a high correlation of 0.85. To the best of our knowledge, our study is the first attempt in applying perceptual models of standard design principles in computational aesthetic evaluation of logos. This work provides a machine-learning scheme for quantitative aesthetic evaluation of graphic design patterns and also enhances understanding of aesthetic perception of logos.

The rest of this article is organized as follows. Section 2 summarizes related work on computational aesthetics. Section 3 introduces the framework of computational aesthetic evaluation of logos. We then evaluate logos on balance (Section 4), contrast (Section 5), and harmony (Section 6) computationally. Section 7 introduces human evaluation survey from which scores given by subjects are used for regression modeling in our experiment. Section 8 presents machine-learning-based regression modeling, and associated model performance analysis is introduced in Section 9, which verify our computational evaluation metrics. Finally, we conclude and discuss our future work in Section 10.

2. RELATED WORK

Computational aesthetics has become an active research field in recent years. Many works mainly focus on data-driven approaches for aesthetic evaluation of artworks, especially paintings and photographs. Several works are focus on aesthetics of web pages. Such works have focused on designing appropriate visual features or rules that attempt to capture specific aesthetic principles and use standard machine-learning techniques such as linear classifiers or regressors to predict aesthetic perception of humans.

There has been research on estimating the aesthetic quality of photographs, including methods to differentiate between photos captured by professional photographers versus amateurs [Dhar et al. 2011; Ke et al. 2006; Obrador et al. 2012; Wong and Low 2009]. Ke et al. [2006] evaluated the aesthetic value of any photo through visual features capturing the spatial distribution of edges, color, blur, and brightness. Wong and Low [2009] used a visual saliency model to classify each image as professional or snapshots. Dhar et al. [2011] used the human-describable attributes, such as composition, illumination, and the image content in the aesthetic model of photos. Obrador et al. [2012] proposed a category-based approach to judge the aesthetic appeal of photographs, in which they computed several features such as simplicity, rule-of-thirds, visual balance, and region contrast. Obrador et al. [2012] also calculated the contrast in sharpness, exposure, saturation, hue, blurring, texture, and edge spatial distribution between contrasting regions. Most visual features abstracted from photographs are related to texture, colorfulness, and brightness, which are absent in black-and-white logos. Most of the composition rules in photographs also differ from the design principles in logos.

For paintings, we mainly refer to the work of Li and Chen [2009], where the authors evaluated the aesthetic quality of paintings by detecting global color distribution, brightness, blur effects, and edge distribution, together with local features such as color, brightness, and blurring contrast in the segmented regions and also the colors of focused regions based on the rule of thirds. Additionally, they calculated color harmony by fitting the hue and saturation-lightness models for oil paintings. A state-of-the-art recognition system to learn the emotional of abstract paintings was proposed in Sartori et al. [2015] based on statistical analysis and art theory. Since logos are usually designed with simple forms with just a few colors or even a single color [Wang 2009], we focus on shape-related features in logos.

On web page aesthetics, the closest related representative works are those of Zheng et al. [2009] and Reinecke et al. [2013]. Zheng et al. [2009] used low-level image statistics, such as symmetry, balance, and equilibrium, to analyze the layout structure of a website. Reinecke et al. [2013] extended Zheng's work by providing the first perceptual models of visual complexity and colorfulness in websites and used these two models to predict user's first aesthetic impressions of websites. The aesthetic factors in web pages, including colorfulness, space-based decomposition, symmetry, and balance based on quad-tree decomposition are closely related to web page layout characteristics. Since logos differ considerably from websites, the evaluation frameworks of web page aesthetics are also unsuited to logos.

Liao and Chen [2014] incorporated visual feature extraction and analysis algorithms commonly utilized in computer vision to compute proper indices and investigate key visual elements in logo designs, including complexity, balance, and repetition. They performed statistical analysis on the distribution of computed indices of complexity, balance, and repetition but did not conduct human evaluation to validate their metrics. Wang [2009] found out the color regularity in enterprise logos by calculating the number of colors and hues, as well as the relevance of Hue, Saturation, Brightness (HSB) data. Rauschenberger et al. [2009] used the predominant spatial frequency of icons for efficient subset search. Li et al. [2014] proposed the automatic generation of logo designs by extracting the shape and color grammar in a logo's visual structure. With crowdsourced data on the human perception of similarity, O'Donovan et al. [2014] proposed interfaces for font selection based on estimating attribute values and similarity of visual fonts. Garces et al. [2014] presented a style distance function for measuring style similarity for vector art. Saleh et al. [2015] modeled stylistic similarity for infographics using low-level visual features. Recently, a new method was presented by Laursen et al. [2016] that selects a complete icon set optimized for comprehensibility and identifiability, given several icon candidates representing functionality.

Although unrelated with computational aesthetic evaluation of logo designs, other works build a bridge between aesthetic perception of human and computer vision. Thumfart et al. [2011] modeled the hierarchical relationship between human aesthetic texture perception and computational texture features by building a layered model. Wallraven et al. [2009] used an eye tracer to investigate eye movements when people appreciate artworks, in an attempt to build a model of human aesthetics. Henderson and Cote [1998] proposed guidelines and related characteristics to select or modify logos including activity, balance, symmetry, harmony, proportion, repetition, and so on. The beauty of visual objects had been shown to be affected by factors such as symmetry, contrast, complexity, and perceptual fluency in Reber et al. [2004]. Although these works appear not directly related to our study, they provide inspirations on selecting important visual features and aesthetic rules that closely related to aesthetic perception of human and also towards logo designs. Our framework of computational aesthetic evaluation of logos builds relationship between design principles and aesthetic perception of logos, as detailed in the following sections.
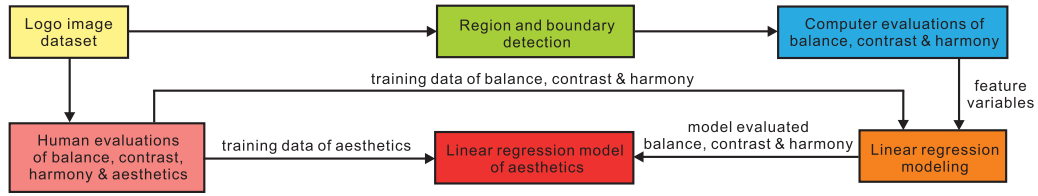
Fig. 2.   Overview of the framework of our computational aesthetic evaluation of logos. The input logo image dataset (yellow box) is first used for extracting feature variables (green and blue boxes) on the one hand and then collecting human evaluation scores of balance, contrast, harmony, and aesthetics (pink box) on the other hand. Then we regress human-evaluated scores on those feature variables to obtain linear regression models of balance, contrast, and harmony (orange box), which are finally used as new input variables to obtain linear regression model of aesthetics (red box).

3.   A FRAMEWORK OF COMPUTATIONAL AESTHETIC EVALUATION OF LOGOS

Figure 2 shows our framework of computational aesthetic evaluation of logos. For each logo image in the image dataset (yellow box), we proceed in four steps. First, we detect regions and associated boundaries of each logo (green box). Second, we construct several feature variables and use them to evaluate the balance, contrast, and harmony of each logo (blue box). To validate the effectiveness of our computer evaluation metrics, we conduct human evaluations of balance, contrast, harmony, and aesthetics (pink box). Third, we regress human-evaluated balance, contrast, and harmony scores on those feature variables to obtain corresponding linear regression models (orange box). Finally, we regress human-evaluated aesthetics scores on model-evaluated balance, contrast, and harmony to obtain the final linear regression model of aesthetics (red box).

A logo, denoted as G, may have one or more regions. These regions are jointed and organized by design principles such as balance, contrast, harmony, variety, and rhythm to form a complete logo, thus, they may exhibit some "holes" of varying shapes inside G, as shown in Figure 1. We call each region of G a visual element. Here we use the regionprops function in Matlab to detect the regions in G and denote each region as E_k, $(k = 1, 2, \ldots, N$, where N is the number of regions detected in G). Then we detect the boundary B_k of E_k using the Canny edge detection algorithm, and the corresponding boundary points are put into a set denoted as BP_k. Note that B_k may represent both the exterior boundaries of G and the boundaries of interior holes inside G.

The following three sections proceed to detail the computational aesthetic evaluation of logos on balance, contrast, and harmony. For each dimension, we will present the metrics of feature variable extraction and evaluation.

4.   BALANCE

Balance refers to the equal distribution of visual weight in a work of art; it strives for a state of equilibrium to create a sense of tranquility. The most commonly used balance in logos is the left-right balance, and logos with left-right symmetrical balance [Liao and Chen 2014] appear stable (the leftmost column in Figure 1), while those logos appearing more dynamic are designed with asymmetrical balance (the rightmost column in Figure 1). Thus, it is desirable to measure the degree of left-right balance in G, which is more challenging than simply detecting left-right symmetry. Here we construct a *weight difference curve* to measure the degree of left-right balance in G.

We denote the height and width of G as H and W, respectively, the vertical symmetric axis of G is denoted as X_V. Consider that the white regions reflect the space left by the black regions that together form the logo's shape (see Figure 1), and G is balanced if the total weight of black regions on the left side and right side of X_V are equal. Thus, we scan G from top to bottom and line by line. On each line

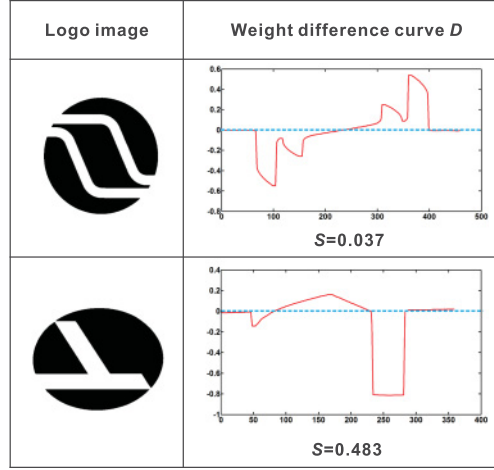| Logo image	Weight difference curve D
	$S=0.037$
	$S=0.483$

Fig. 3. Weight difference curves D (in red) associated with two logos. The horizontal axis (blue dashed lines) represents the index of the scanned line, and the vertical axis represents the weight difference on the corresponding scanned line. S is calculated as the square root of the absolute difference between the area covered by D above (positive) and below (negative) the horizontal axis. The top logo has smaller S (close to zero) and higher degree of left-right balance than the bottom logo.

scanned, we first pick a pixel $P_{i,j}$ falling into the black regions in G from left to right ($i = 1, 2, \ldots H$; $j = 1, 2, \ldots W_i$, where i is the index of the scanned lines, j is the index of pixels picked on the ith scanned line, and W_i is the number of pixels picked on the ith scanned line).

We define dl_i and dr_i as the weight set of the black regions on the left side and right side of X_V on the ith scanned line, respectively. Then we calculate the absolute distance $d_{i,j}$ between $P_{i,j}$ and X_V and put its normalized form (dividing $d_{i,j}$ by $0.5 \cdot W$) into dl_i if $P_{i,j}$ is on the left side of X_V and otherwise dr_i if $P_{i,j}$ is on the right side of X_V. The weight difference between dl_i and dr_i on the ith scanned line is measured by D_i defined as

$$D_i = \frac{1}{Ml_i} \sum_{j=1}^{Ml_i} dl_{i,j} - \frac{1}{Mr_i} \sum_{k=1}^{Mr_i} dr_{i,k}, \tag{1}$$

where Ml_i and Mr_i are the number of elements in dl_i and dr_i, respectively, $dl_{i,j}$ is the jth element in dl_i, as the same for $dr_{i,k}$ in dr_i. We calculate D_i of all the scanned line for $i \in [1, H]$ and obtain the weight difference curve D. Two examples of D associated with two logos are shown in Figure 3, where the horizontal axis (blue dashed lines in Figure 3) represents ith scanned line and the vertical axis represents D_i. Obviously, the whole G is recognized as being left-right symmetrical balanced if $D_i = 0$ for $i \in [1, H]$. Otherwise, the logo G must contains some tilt components with varying zigzag tendencies from top to bottom. Here we denote S as the square root of the absolute difference between the area covered by D above (positive) and below (negative) the horizontal axis and use it to measure the degree of left-right balance in G. The smaller the S the higher degree of left-right balance in G. We will validate S through regression analysis in Section 8 and obtain the final prediction model for perceived balance.

5. CONTRAST

Contrast refers to differences in values (light and dark values), colors, textures, shapes, and other elements. Contrast creates visual excitement and adds interest to the design. Since existing logos could have multiple colors of multiple values, a single color of a single value (such as MacDonald), or
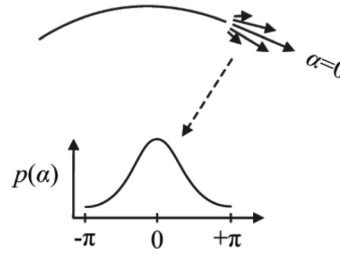
Fig. 4. Von Mises distribution. The expected change in tangent direction α is distributed as a von Mises distribution centered on 0 (straight).

no color (such as those metal auto logos), we converted all the collected logo images into black and white ones so we can focus on the shape contrast in this study.

Bar and Neta [2006] found that the sharp corners on objects seen in our environment activate fear processing in the brain amygdala area. Thus shapes with corners draw more attention than rounded shapes. They further indicated that humans are born with a deeply rooted preference for curves and prejudice against sharp angles. In design, corners help attract visual attention and curved items help create a positive mood and aesthetic impression. We therefore propose *hard-soft contrast* to evaluate the shape contrast in logos.

In an experiment, Attneave found that information is concentrated along contours of shapes and is further concentrated at those points on a contour at which its direction changes most rapidly [Attneave 1954]. Motivated by this finding, we adopt Shannon's information theory [Shannon 1948] to evaluate the degree of hard-soft contrast. Our idea is to choose the curvature of BP_k (Section 3) as a random variable and calculate the information associated with corner points (hard parts) as well as sampled points (soft parts) in G; the strength of the hard-soft contrast is then measured by the ratio of information associated with hard parts to that associated with both hard and soft parts, as detailed next.

It has been shown [Page et al. 2003] that, in the discrete case, when a curve is uniformly sampled, the curvature is directly proportional to the turning angle α (i.e., change of tangent direction from point to point along the shape contour). We assume that it follows a von Mises distribution centered on "straight" $\alpha = 0$ [Feldman and Singh 2005],

$$p(\alpha) = A \, \exp\left[b cos(\alpha)\right], \tag{2}$$

where b is a parameter modulating the spread of the distribution and A is a normalizing constant (depending on b but not α) and positive values of α correspond to clockwise turns and negative values to counterclockwise turns (see Figure 4). We select $A = 0.36$ and $b = 1$ in our experiment.

We define Ψ_c and Ψ_s as the corner points set and sampled points set of G, respectively. Since the turning angle α and the vertex angle θ satisfy $\alpha + \theta = \pi$, we can obtain α with $\alpha = \pi - \theta$ if θ is estimated. To obtain more appropriate values of vertex angles, for each visual element E_k (Section 3) in G, we first detect its corner points using the *CSS* corner detector [He and Yung 2004] and put them into Ψ_c, as indicated by the red points in Figure 5. Those corner points correspond to the sharp vertices in BP_k and divide BP_k into several pieces, and those pieces correspond to the rounded or straight parts in BP_k. Next, on each piece with two corner points on its two ends, we start from one corner point and sample the piece every three points along BP_k as *break points*. Finally, we take the corner points and break points as control points and apply the arc length parametrization on the chosen piece so uniform sampling can be achieved [Feldman and Singh 2005]. The sampled points are put into Ψ_s (except the corner points), as indicated by the green points in Figure 5. Besides, the sampled points (including the
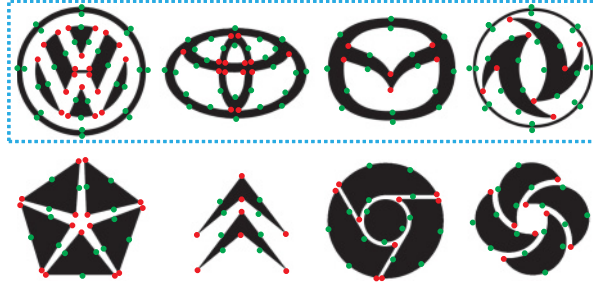
Fig. 5. Some corner (in red) and sampled points (in green) detected in logos. The corner points correspond to the sharp vertices (hard parts), which divide each boundary into several round or straight pieces. The sample points are taken on each piece, which represent the soft parts in logos. The logos on the top row (blue dashed box) have both exterior and interior boundaries, and those on the bottom row have no closed exterior boundaries.

corner points) are put into a feature points set denoted as FP_k, which is used in Section 6. We continue this procedure until all pieces in BP_k are sampled.

Shannon defined the information as the negative of the logarithm of the probability distribution [Shannon 1948]. Thus, the information I_α associated with a vertex (chosen from either Ψ_c or Ψ_s) having a turning angle α (denoted as $vertex(\alpha)$) is calculated by

$$I_\alpha = -\log_2 p(\alpha). \tag{3}$$

The information associated with hard parts (corner points) and soft parts (sampled points) in G are calculated by

$$I_h = -\sum_{vertex(\alpha)\in \Psi_c} \log_2 p(\alpha) \tag{4}$$

and

$$I_s = -\sum_{vertex(\alpha)\in \Psi_s} \log_2 p(\alpha), \tag{5}$$

respectively. Finally, the strength of the hard-soft contrast C_{hs} in G is measured by

$$C_{hs} = \frac{I_s}{I_h + I_s}, \tag{6}$$

where C_{hs} is in the interval $[0, 1]$. The logo G appears softer when C_{hs} is close to 1 and sharper if it is close to 0. Initially, we obtain values of C_{hs} for the given logo samples by Equation (6). Some of the values, however, differ significantly from those -evaluated by humans (Section 7), especially for those logos encircled by one or more closed exterior boundaries, as indicated by those logos on the upper row in Figure 5. We speculate that both exterior boundaries and other interior boundaries (the boundaries of "holes" inside G) contribute independently to the overall strength of the hard-soft contrast in those logos. Thus, we calculate the strength of the hard-soft contrast for both exterior boundaries (if it exists in G) and other interior boundaries, denoted as C_{out} and C_{in}, which are measured by

$$C_{out} = \frac{I_{sout}}{I_{hout} + I_{sout}} \tag{7}$$

and

$$C_{in} = \frac{I_{sin}}{I_{hin} + I_{sin}}, \tag{8}$$
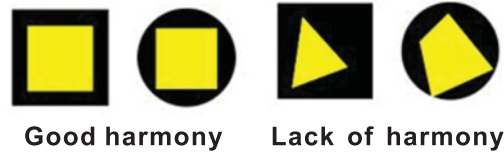
Good harmony   Lack of harmony

Fig. 6.   Harmony in graphic design. The graphic designs on the left are more harmonic than those on the right.
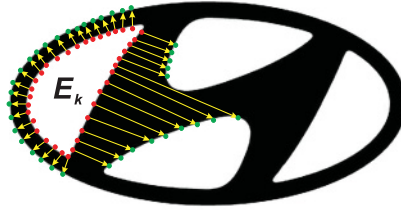

Fig. 7.   Illustration of *normal tracking distance* for a visual element E_k in a logo. We start from each selected point *sp* (in red) and move the point in the outer normal direction (yellow arrow) until it reaches the boundary of another adjoining visual element at the destination point *dp* (in green), the distance between *sp* and *dp* is called *normal tracking distance*.

respectively. Here I_{hout} and I_{sout} represent the information associated with hard parts and soft parts in exterior boundaries, as for I_{hin} and I_{sin} in interior boundaries. For those logos without closed exterior boundaries, as indicated by those on the bottom row in Figure 5, we simply calculate C_{hs} in G to obtain the overall strength of the hard-soft contrast in G, where C_{out} and C_{in} are equal to C_{hs}. We will regress the human-evaluated hard-soft contrast on C_{out} and C_{in} to obtain the final prediction model for the perceived contrast, to be described in Section 8.

6.   HARMONY

Harmony is a measure on how the visual elements in an art work fit together or how they belong together. It provides the cohesive quality that makes an art work complete. For visual elements in a logo design, harmony implies that the shape of one visual element should fit the shape of its adjoining visual elements, as illustrated in Figure 6. In addition to individual visual elements, many logos may exhibit "holes" (the white regions in Figure 7) of varying shapes inside the exterior boundaries. Those holes thus become individual visual elements and the region between adjoining holes (the black region in Figure 7) would also contribute to the overall harmony of logos, as shown in Figure 7. In this section, we define *normal tracking distance* to calculate the goodness of fit between each adjoining visual elements in a logo, as detailed next.

Since we have detected the boundary B_k of each visual element E_k (Section 3) in G, as well as its feature points set FP_k (Section 5), we connect every two feature points by a line segment so B_k can be approximated by the line segments. Next, we select a point on B_k that is nearest to the middle point of a segment, indicated as the red points in Figure 7. Denoting this point as *sp*, we start from *sp* and move a destination point in the outer normal direction of the line segment until it reaches the boundary of another adjoining visual element, the reached point is denoted as *dp*, indicated as the green points in Figure 7. We then connect from *sp* to *dp* to form an arrow (yellow in Figure 7). We call the distance between *sp* and *dp* as *normal tracking distance* (denoted as *ntd*). The distance is put into a distance set denoted as ntd_k. When all feature points in FP_k are visited, we calculate the variance v_k of all the distances in ntd_k to measure the goodness of fit between E_k and its adjoining visual elements.

Fig. 8. Experimental image dataset. All logo images are selected from the famous autos and computer brands, as well as aviation companies.
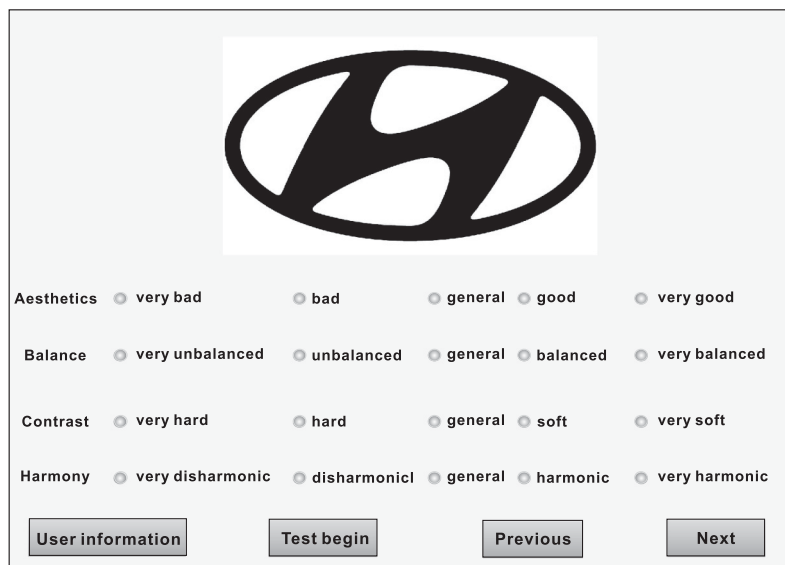
Note that v_k is in the interval $[0, 1]$, a value of v_k close to 0 indicates that E_k is separated from its adjoining visual elements with nearly equidistance, and thus E_k and its adjoining visual elements together look more harmonic. While a value of v_k close to 1 indicates that there are mutations between E_k and its adjoining visual elements, as a result, they look less harmonic. We use the average value of v_k over all visual elements in G, denoted as Fit, to measure the overall harmony of a logo. We will validate Fit in Section 8 through regression analysis and get the final prediction model for the perceived harmony.

7. HUMAN EVALUATION

Being treated as a data-based learning problem, the validation of our aforementioned computational evaluation metrics highly rely on the training data used for learning. Unlike those works on photographs, it is hard to find a website of logos with ratings by a large community. The prevalence of logo designs raises the need of evaluation. Therefore we conduct human ratings on the balance, contrast, harmony, as well as aesthetics of 60 logo images selected from well-known auto, computer, and aviation companies, as shown in Figure 8. The logo images are downloaded as JPEG files through "Google image search" with careful selection on size and definition. We convert the original logo images into black and white, so we are able to concentrate ratings on a few important concepts of logo design, which are also relatively easy to be evaluated computationally.

7.1 Process of Human Evaluation

60 subjects (30 male and 30 female, aged 18–27, including undergraduate and graduate students from the Faculties of Science, Engineering, Medicine and Art) volunteered to participate in rating each aspect of the 60 logos. They are psychologically mature enough to have common aesthetic experiences learned eithr in school or at universities or from their parents, relatives, and friends, so they are qualified to evaluate logos selected from well-known auto, computer, and aviation companies.

Fig. 9. An example page of the survey. Each logo image is displayed in a window above, the subject is required to rating the 4 concepts: "Aesthetics," "Balance," "Contrast," and "Harmony." Each concept has 5 rating options ranging from 1 (leftmost) to 5 (rightmost). After rating the current logo image, the subject clicks on the "Next" button and continues ratings on the next logo image or the "Previous" button to redo the ratings of the previous logo image.

Each logo image has an independent rating page, as shown in Figure 9, where the logo image is displayed in a window of 700×900 pixels. Each subject is required to give four ratings for evaluating the following four concepts of each logo image: "Aesthetics," "Balance," "Contrast," and "Harmony," which are described as below:

Aesthetics Your overall impression on the displayed logo at the first sight.
Balance Your feeling on the distribution of left-right visual weight in the displayed logo.
Contrast Your feeling on the hard-soft contrast in the displayed logo.
Harmony Your feeling on the goodness of fit between adjoining visual elements in the displayed logo.

Here we adopt the most widely used and well-established 5-point Likert scale for rating, ranging from 1 to 5, where 5 (rightmost) means the most positive value and 1 (leftmost) means the the most negative value. Before starting the rating, we explained the meaning of the four concepts to all subjects, so they could focus more on the measurement of the four concepts defined in our work. All subjects received clear instructions about the rating dimensions and tasks. They were also given practice trials to ensure they had a clear understanding before starting the actual experiment. After rating the current logo image, the subject clicked on the "Next" button and continued ratings on the next logo image. If the subject hesitated about the decision made, then he/she could click on the "Previous" button, and the page redisplayed the previous logo image and the subject could redo the ratings. The rating survey terminated when all 60 logo images had been rated. After rating all subjects were asked whether they understood the meaning of each concept to check if they had any further comments or questions. None of the 60 subjects had any trouble in providing their evaluation scores, which ensures that there is no random answer or uniformly neutral rating.
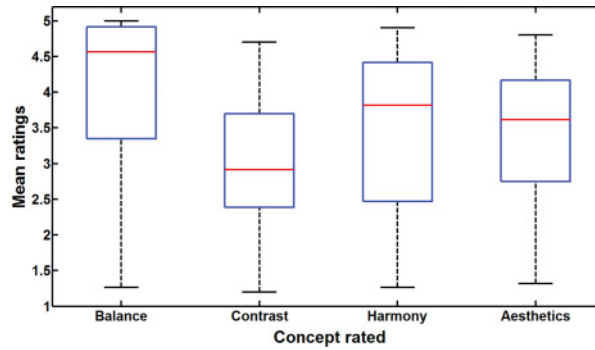
Fig. 10.　Distribution of mean ratings across 60 logos on balance, contrast, harmony, and aesthetics. Each central blue rectangle spanned from the first to third quartile indicates the likely range of variation, the red segment inside the rectangle shows the median, and the "whiskers" above and below the rectangle show the minimum and maximum.

7.2　Statistical Analysis of Human Ratings

We show the mean (denoted as μ) and standard deviation (denoted as σ) of ratings across all subjects for each logo on balance, contrast, harmony, and aesthetics in Table VI. The standard deviations in Table VI show that the rating variances across all subjects are quite small (most of them are less than 0.7). The highly reliable and consistent ratings across all subjects for the 60 logos demonstrates that they all understood the tasks, so they gave their ratings reliably rather than randomly.

We also show the distribution of mean ratings across 60 logos on balance, contrast, harmony, and aesthetics in Figure 10 by using boxplot, which is a standardized way of displaying the distribution of data based on the five number summaries: minimum, first quartile, median, third quartile, and maximum. Each central blue rectangle spans from the first to third quartile (the IQR). A segment (red line) inside the rectangle shows the median, and the "whiskers" above and below the rectangle show the minimum and maximum. The figure displays the full range of variation (from minimum to maximum), the likely range of variation (the IQR), and a typical value (the median). The mean ratings of 60 logos are used as the manually evaluated scores for regression modeling of balance, contrast, harmony, and aesthetics in Section 8.

8.　REGRESSION MODELING

In this section, we use a machine-learning-based linear regression model to analyze the relationship between manually judged balance, hard-soft contrast, harmony, and various feature variables extracted from geometric properties of logo images in the previous sections. We also analyze the relationship between manually judged aesthetics and balance, hard-soft contrast, and harmony evaluated by our regression models. Here we adopt statistical analysis software SPSS [Bryman and Cramer 2011] to perform the regression on balance, hard-soft contrast, harmony, and aesthetics, because it offers a full package of statistics used in regression analysis. Regression analysis allows us to assess the validity of our regression models in terms of goodness of fit, check statistical significance by F-test of the overall fit, and use t-test to select predictor variables that are significant.

To simplify comparison of the relative importance of each variable, we normalize each variable by feature scaling. Then, for each dimension of balance, hard-soft contrast, and harmony, we adopt the method of Monte Carlo cross-validation [Dubitzky et al. 2007] during the model training. This is an effective and well-accepted model validation technique in machine learning to avoid overfitting by determining how well our model generalizes to an independent dataset. The method randomly splits

Table I. Statistics of Regression Model for Perceived Balance C_b

R^2=.943	$Adj.R^2$=.941		$Std.Er$=.267	F=470.736	$Sig.$=.000
Var	$Reg.Coef$	$Std.Er$	$Std.Coef$	t	$Sig.$
S	.996	.051	.971	13.908	.000
(Const.)	−.155	.042		−3.548	.003

Table II. Statistics of Regression Model for Perceived Hard-Soft Contrast C_c

R^2=.935	$Adj.R^2$=.930		$Std.Er$=.257	F=226.206	$Sig.$=.000
Var	$Reg.Coef$	$Std.Er$	$Std.Coef$	t	$Sig.$
C_{out}	.299	.038	.451	7.914	.000
C_{in}	.527	.047	.648	10.862	.000
(Const.)	.388	.098		3.968	.000

Table III. Statistics of Regression Model for Perceived Harmony C_h

R^2=.866	$Adj.R^2$=.863		$Std.Er$=.301	F=176.968	$Sig.$=.000
Var	$Reg.Coef$	$Std.Er$	$Std.Coef$	t	$Sig.$
Fit	.824	.055	.945	14.266	.000
(Const.)	.437	.148		2.872	.006

the dataset into the training set and testing set (each set contains 30 logo images). For each split, the model is fit to the training set and validated using the testing set. We repeat this process for eight runs for each dimension, and the results are then averaged over eight splits to get the final regression model of each dimension, as shown in Tables I (Section 8.1), II (Section 8.2), and III (Section 8.3), respectively. Next, we take these three models as predictors to obtain the regression model of aesthetics based on the method proposed in Reinecke et al. [2013], and Table IV (Section 8.4) shows statistics of our regression model for perceived aesthetics of logos. The statistical meanings associated with Tables I– IV are briefly explained in Section 8.5.

8.1 Computer-Evaluated Balance

The statistics of the regression model for perceived balance is shown in Table I. S is the square root of the absolute area covered by D (obtained with Equation (1)) and the horizontal axis. Table I indicates that our model is able to explain 94.1% of the variance in human-perceived balance and is statistically significant ($Sig. = .000$). S makes a significant contribution to our model, since its associated p-value is less than 0.05. These show that the weight difference curve D calculated in Section 4 can properly describes the distribution of left-right visual weight in logos and S can efficiently describe the degree of perceived balance. Our equation of regression model for perceived balance is

$$C_b = -.155 + .996 \cdot S. \tag{9}$$

8.2 Computer-Evaluated Hard-Soft Contrast

The statistics of regression model for perceived contrast is shown in Table II. C_{out} and C_{in} describe the strength of the hard-soft contrast for the exterior and interior boundaries in logos with Equations (7) and (8), respectively. Table II indicates that our model is able to explain 93% of the variance in human-perceived hard-soft contrast and is statistically significant ($Sig. = .000$). C_{out} and C_{in} make significant contributions to our model, since their associated p-values are less than 0.05. The coefficients of C_{out}

Table IV. Statistics of Regression Model for Perceived Aesthetics C_a

$R^2{=}.985$	$Adj.R^2{=}.983$		$Std.Er{=}.321$	$F{=}998.256$	$Sig.{=}.000$
Var	$Reg.Coef$	$Std.Er$	$Std.Coef$	t	$Sig.$
C_h	.771	.045	.788	16.102	.000
C_c	.172	.055	.146	5.807	.000
C_b	.086	.046	.073	2.903	.000

and C_{in} ($Reg.Coef$ column) indicate that C_{in} contributes slightly more than C_{out} in the human-perceived hard-soft contrast of logos. Our equation of regression model for perceived hard-soft contrast is

$$C_c = .388 + .299 \cdot C_{out} + .527 \cdot C_{in}. \tag{10}$$

8.3 Computer-Evaluated Harmony

The statistics of regression model for perceived harmony is shown in Table III, and Fit measures the average goodness of fit between all adjoining visual elements in logos. The model is able to explain 86.3% of the variance in human-perceived harmony and is statistically significant ($Sig. = .000$). Fit makes a significant contribution to our model, since its associated p-value is less than 0.05. These show that the *normal tracking distance* calculated in Section 6 can properly describe the goodness of fit between all adjoining elements of a logo and Fit can efficiently describes the degree of perceived harmony. Our equation of regression model for perceived harmony is

$$C_h = .437 + .824 \cdot Fit. \tag{11}$$

8.4 Computer-Evaluated Aesthetics of Logos

We also build a regression model for human-perceived aesthetics of logos by taking C_b, C_c, and C_h as predictor variables, and the overall statistics of the regression model is shown in Table IV. The model measured by C_h, C_c, and C_b is able to explain 98.3% of the variance in human-perceived aesthetics of logos and is statistically significant ($Sig. = .000$). Here C_h, C_c, and C_b all make significant positive effects on our model, since their associated p-values are less than 0.05. The coefficients of C_h, C_c, and C_b ($Reg.Coef$ column) indicate that the harmony C_h plays a more important role than the hard-soft contrast C_c and balance C_b in predicting the human-perceived aesthetics of logos. Our equation of regression model for perceived aesthetics is

$$C_a = .771 \cdot C_h + .172 \cdot C_c + .086 \cdot C_b. \tag{12}$$

8.5 Regression Statistics

The statistics of overall model fit are given on the first row of Tables I–IV.

- The coefficient of determination, R^2, is the proportion of variance in the dependent variable that can be explained by the independent variables. Statistically, any regressor with R^2 larger than 0.8 is regarded as strong and less than 0.5 as weak.
- The adjusted R^2 ($Adj.R^2$) is an adjustment of R^2 that penalizes the addition of extraneous predictors to the model.
- Standard Error ($Std.Er$) of the estimate is the average distance between observed values and regression line. The smaller the value is, the more accurate the predictions are.
- F and $Sig.$ are, respectively, the F-statistic and associated p-value. F is the mean square of regression divided by the mean square of residual, and $Sig.$ indicates the statistical significance of the regression model. If the p-value is less than or equal to 0.05, then the regression model statistically significantly predicts the outcome variable.

(a) Balance   (b) Contrast

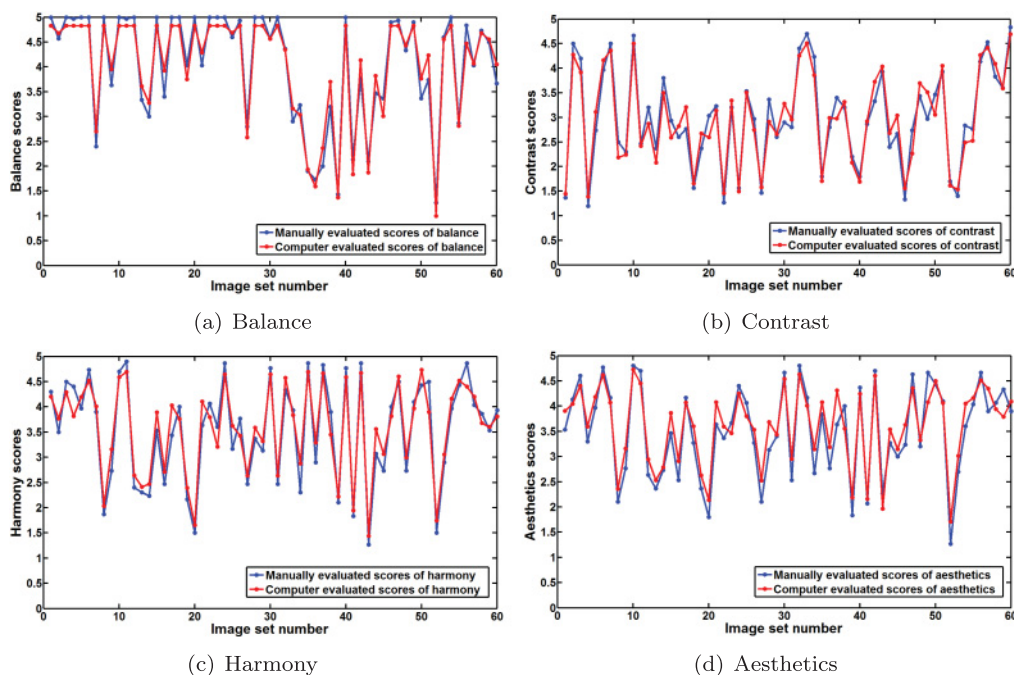(c) Harmony   (d) Aesthetics

Fig. 11. Ranking scores of 60 logos evaluated by subjects (in blue) and regression models (in red) in the image dataset.

The remaining rows in Tables I– IV are parameter estimates.

- The *Var* column shows predictor variables (where Const. stands for Constant).
- Regression Coefficients (*Reg.Coef*) are the coefficients for the regression equation.
- Standard Error (*Std.Er*) are the standard errors associated with the coefficients. The smaller the values are, the more accurate the model estimates the coefficients.
- Standard Coefficients (*Std.Coef*) are the coefficients obtained after standardizing all of the variables in the regression, including all of the dependent and independent variables, and running the regression.
- *t* and *Sig.* are, respectively, the T-statistics and their associated two-tailed *p*-values used for testing whether each coefficient significantly differs from zero at the level of 0.05.

9. ANALYSIS OF OUR REGRESSION MODELS

9.1 Comparison between Computer and Human Evaluations

We have obtained regression models of balance, hard-soft contrast, harmony, and aesthetics of logos in Section 8. Figure 11 shows the average ranking scores of 60 logos evaluated across all subjects in Section 7 (blue lines) and ranking scores evaluated by regression models (red lines).

To evaluate the prediction performances of our regression models, we show in Table V the average Pearson correlation coefficient and associated two-tailed Significance (denoted as Average R_P/$Sig.$) and average mean squared error (denoted as Average MSE) between model-evaluated scores using C_b, C_c, C_h, and C_a in Section 8 and manually evaluated scores of each dimension in Section 7 over 8 runs for both training and testing sets to evaluate the average prediction performances of our regression models for different training and testing sets. We also show the standard deviation (denoted as SD) of

Table V. Statistics of Performances Associated with Different Models over Eight Runs, Where S_1 is the Training Set and S_2 Is the Testing Set

	C_b	C_c	C_h	C_a
Average $R_P/Sig.$ (S_1)	.978/.000	.947/.000	.883/.000	.861/.000
Average $R_P/Sig.$ (S_2)	.964/.000	.924/.000	.878/.000	.851/.000
SD of $R_P/Sig.$ (S_1)	.011/.000	.023/.000	.014/.000	.016/.000
SD of $R_P/Sig.$ (S_2)	.012/.000	.020/.000	.015/.000	.010/.000
Average MSE (S_1)	.054	.074	.095	.110
Average MSE (S_2)	.076	.088	.105	.123
SD of MSE (S_1)	.011	.014	.016	.012
SD of MSE (S_2)	.012	.009	.010	.007

(a) R_P correlation coefficient   (b) MSE
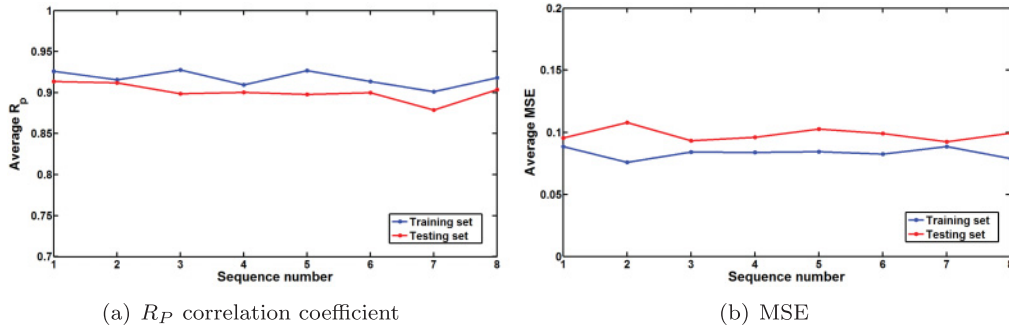
Fig. 12. Performance of training (in blue) and testing set (in red) with (a) high average R_P correlation coefficient (above 0.85) and (b) low average MSE (below 0.15) over regression models of balance, hard-soft contrast, harmony, and aesthetics for eight runs.

$R_P/Sig.$ and MSE over 8 runs for both training and testing sets to evaluate the variation of prediction performances in Table V. The table clearly shows that our regression models yield a highly significant correlation between manually and computer-evaluated results with high accuracy, since the average R_P are high and associated average $Sig.$ are less than 0.05 with low average MSE, besides the low SD of $R_P/Sig.$ and MSE indicate that the variation of prediction performances is small over eight runs. Our regression models have high prediction performances and the explanatory variables in regression models have high predictive power in computational aesthetic evaluation of logos.

Additionally we perform analysis of variance (ANOVA) and the F/p between model-evaluated scores and manually evaluated scores of 60 logos on balance, hard-soft contrast, harmony, and aesthetics are .019/.891, .231/.633, .349/.557, and .753/.393, respectively. The p value is larger than 0.05 for each model, which indicates that there is no statistically significant difference between model-evaluated scores and manually evaluated scores.

9.2 R_P and MSE over Regression Models for Eight Runs

Figure 12 shows the average R_P and average MSE over regression models of balance, hard-soft contrast, harmony, and aesthetics for eight runs, where blue lines represent the training set and red lines the testing set. We can observe that the regression models have high performances with high average R_P and low average MSE for both training and testing sets.

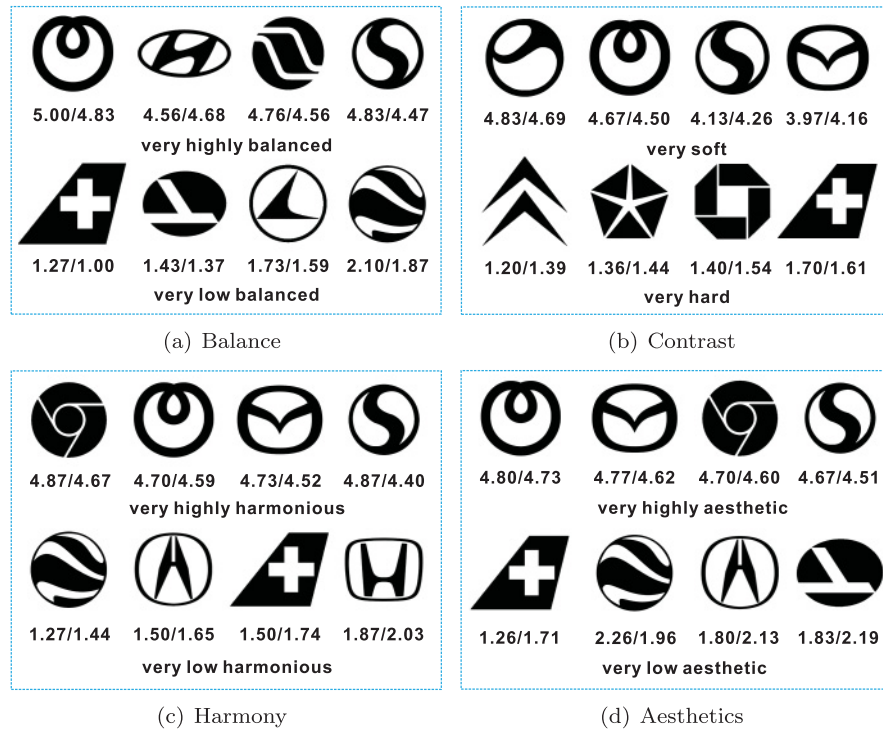|	
5.00/4.83 4.56/4.68 4.76/4.56 4.83/4.47	4.83/4.69 4.67/4.50 4.13/4.26 3.97/4.16
very highly balanced	very soft
1.27/1.00 1.43/1.37 1.73/1.59 2.10/1.87	1.20/1.39 1.36/1.44 1.40/1.54 1.70/1.61
very low balanced	very hard
(a) Balance	(b) Contrast
4.87/4.67 4.70/4.59 4.73/4.52 4.87/4.40	4.80/4.73 4.77/4.62 4.70/4.60 4.67/4.51
very highly harmonious	very highly aesthetic
1.27/1.44 1.50/1.65 1.50/1.74 1.87/2.03	1.26/1.71 2.26/1.96 1.80/2.13 1.83/2.19
very low harmonious	very low aesthetic
(c) Harmony	(d) Aesthetics

Fig. 13. Examples of logos and associated scores evaluated in opposite positions by subjects (in front of the slash) and regression models (at the back of the slash) of all predicted attributes. The logos on the top row in each blue box are scored very high and those on the bottom row are scored very low.

9.3 Examples of Logos Associated with Scores Evaluated in Opposite Positions

Further, we show some examples of logos associated with scores evaluated in opposite positions for each predicted attribute by subjects and regression models of balance, hard-soft contrast, harmony, and aesthetics, respectively, in Figure 13. The evaluated scores of logos on the top row are very high and those on the bottom row are very low. The figure indicates that our regression models can describe each predicted attribute of logos in accord with human perception.

10. CONCLUSIONS AND FUTURE WORK

Building a connection between human perception of design principles revealed in logos and computational visual features extracted from logos is a challenging multidisciplinary problem. Our work explores the relationship between human aesthetic perception and design principles. Experimental results show that our model predicted balance, hard-soft contrast, harmony, and aesthetics have highly significant correlation with human evaluations on the same dimensions. To develop efficient feature metrics is a crucial part for evaluating each design principle. Although features extracted here are low level, all of them implicitly describe corresponding definition of each design principle revealed in logos. Our work provides a machine-learning-based reference scheme for quantitative aesthetic evaluation of graphic design patterns. Although our work is not meant to provide a full solution, it is a step forward in this new and interesting research direction.

The methods described in this article have several limitations.

- Currently, we do not consider the color factor, while color is certainly important in aesthetic evaluation of logo images.
- The logos selected in our study are representatives of name-brand companies, professionally designed with possibly high levels of balance, harmony, and aesthetics. Our trained models, therefore, could potentially be generalized to the professional design logos. But they may not work very well with poorly or randomly designed logos.
- Figurative logos, those with representational objects, are not included in our logo image set, since they contain an emotion factor that could potentially affects human aesthetic perception.

In our future research, we will include the color factor in our aesthetic evaluation of logo images and include casually designed logos for training in our models. We will also explore high-level semantic features and extract emotion-related information in the aesthetic evaluation of figurative object-based logo images. Moreover, we will use our work to guide and verify rule-based automatic generation of logo designs [Li et al. 2014].

APPENDIX

A. STATISTICS OF HUMAN RATINGS ACROSS 60 LOGOS

The mean (denoted as μ) and standard deviation (denoted as σ) of ratings across all subjects of each logo on balance, contrast, harmony, and aesthetics are shown in Table VI.

Table VI. Statistics of Human Ratings on Balance, Contrast, Harmony, and Aesthetics across 60 Logos

Logo number	Balance (μ/σ)	Contrast (μ/σ)	Harmony (μ/σ)	Aesthetics (μ/σ)
1	5.00/0.00	1.36/0.49	4.30 /0.47	3.53 /0.63
2	4.56/0.50	4.50/0.57	3.50/0.51	4.13/0.57
3	5.00/0.00	4.20/0.76	4.50/0.57	4.60/0.49
4	4.97/0.18	1.20/0.41	4.40/0.62	3.30/0.75
5	5.00/0.00	2.73/0.45	3.96/0.61	3.96/0.72
6	5.00/0.00	3.97/0.41	4.73/0.45	4.77/0.43
7	2.40/0.49	4.50/0.51	3.90/0.76	4.17/0.79
8	5.00/0.00	2.50/0.51	1.86/0.35	2.10/0.66
9	3.63/0.49	2.30/0.46	2.73/0.58	2.76/0.63
10	5.00/0.00	4.66/0.48	4.70/0.47	4.80/0.41
11	4.97/0.18	2.47/0.63	4.90/0.31	4.70/0.46
12	5.00/0.00	3.20/0.41	2.40/0.56	2.63/0.76
13	3.33/0.61	2.37/0.49	2.30/0.46	2.37/0.61
14	3.00/0.37	3.87/0.73	2.23/0.50	2.73/0.78
15	5.00/0.00	2.93/0.25	3.53/0.51	3.47/0.82
16	3.40/0.50	2.60/0.50	2.46/0.57	2.53/0.68
17	5.00/0.00	2.77/0.68	3.43/0.68	4.17/0.75
18	5.00/0.00	1.57/0.68	4.00/0.37	3.27/0.64
19	4.03/0.67	2.37/0.49	2.17/0.53	2.37/0.61
20	5.00/0.67	3.03/0.18	1.50/0.57	1.80/0.55
21	4.03/0.61	3.23/0.57	3.63/0.72	3.63/0.49
22	5.00/0.00	1.27/0.45	4.07/0.64	3.37/0.61
23	5.00/0.00	3.20/0.66	3.60/0.77	3.67/0.48

(Continued)

Table VI. Continued

Logo number	Balance (μ/σ)	Contrast (μ/σ)	Harmony (μ/σ)	Aesthetics (μ/σ)
24	5.00/0.00	1.57/0.57	4.87/0.34	4.40/0.56
25	4.60/0.50	3.53/0.63	3.17/0.59	4.07/0.64
26	4.93/0.36	2.97/0.18	3.77/0.43	3.27/0.83
27	2.80/0.55	1.47/0.51	2.46/0.51	2.10/0.66
28	5.00/0.00	3.37/0.49	3.37/0.49	3.13/0.73
29	5.00/0.00	2.60/0.50	3.13/0.63	3.40/0.85
30	4.57/0.50	2.90/0.40	4.77/0.43	4.67/0.48
31	5.00/0.00	2.80/0.55	2.46/0.63	2.53/0.51
32	4.36/0.49	4.40/0.62	4.33/0.48	4.80/0.41
33	2.90/0.48	4.70/0.54	3.93/0.25	4.17/0.70
34	3.23/0.63	4.23/0.50	2.30/0.60	2.67/0.61
35	1.90/0.31	1.80/0.66	4.87/0.34	3.83/0.38
36	1.73/0.64	2.80/0.41	2.90/0.76	2.77/0.43
37	2.00/0.64	3.40/0.77	4.83/0.38	3.63/0.57
38	3.20/0.61	3.20/0.41	3.90/0.31	4.00/0.45
39	1.43/0.57	2.20/0.48	2.60/0.72	1.83/0.59
40	5.00/0.00	1.80/0.81	4.77/0.43	4.37/0.49
41	2.13/0.78	2.87/0.68	1.83/0.38	2.07/0.64
42	3.77/0.50	3.33/0.88	4.87/0.35	4.70/0.46
43	2.10/0.48	3.93/0.58	1.27/0.45	2.26/0.69
44	3.47/0.73	2.40/0.62	3.07/0.87	3.27/0.74
45	3.37/0.55	2.67/0.61	2.73/0.74	3.00/0.83
46	4.90/0.55	1.33/0.48	4.00/0.18	3.23/0.90
47	4.93/0.36	2.73/0.58	4.50/0.51	4.63/0.49
48	4.33/0.96	3.43/0.68	2.73/0.69	3.20/0.66
49	4.90/0.55	2.97/0.76	4.10/0.66	4.67/0.55
50	3.37/0.49	3.47/0.63	4.43/0.68	4.43/0.50
51	3.73/0.45	3.93/0.45	4.50/0.63	4.10/0.66
52	1.27/0.45	1.70/0.70	1.50/0.51	1.26/0.45
53	4.60/0.50	1.40/0.63	2.90/0.76	2.70/0.75
54	5.00/0.00	2.83/0.65	3.97/0.18	3.60/0.50
55	2.87/0.51	2.77/1.07	4.43/0.57	4.03/0.72
56	4.83/0.38	4.13/0.51	4.87/0.35	4.67/0.48
57	4.03/0.41	4.53/0.63	4.03/0.89	3.90/0.71
58	4.73/0.45	3.83/0.59	3.87/0.68	4.07/0.69
59	4.50/0.51	3.60/0.62	3.53/0.78	4.33/0.55
60	3.67/0.48	4.83/0.47	3.93/0.25	3.90/0.55

REFERENCES

F. Attneave. 1954. Some informational aspects of visual perception. *Psychol. Rev.* 61, 3 (May. 1954), 183–193. DOI:http://dx.doi.org/10.1037/h0054663

M. Bar and M. Neta. 2006. Humans prefer curved visual objects. *Psychol. Sci.* 17, 8 (Aug. 2006), 645–648. DOI:http://dx.doi.org/10.1111/j.1467-9280.2006.01759.x

A. Bryman and D. Cramer. 2011. *Quantitative Data Analysis with IBM SPSS 17,18 and 19: A Guide for Social Scientists* (1st. ed.). Routledge Press, New York, NY.

S. Dhar, V. Ordonez, and T. L. Berg. 2011. High level describable attributes for predicting aesthetics and interestingness. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'11)*. IEEE, Washington, DC, 1657–1664. DOI:http://dx.doi.org/10.1109/CVPR.2011.5995467

W. Dubitzky, M. Granzow, and D. P. Berrar. 2007. *Fundamentals of Data Mining in Genomics and Proteomics* (1st. ed.). Springer, New York, NY.

J. Feldman and M. Singh. 2005. Information along contours and object boundaries. *Psychol. Rev.* 112, 1 (Jan. 2005), 243–252. DOI:http://dx.doi.org/10.1037/0033-295X.112.1.243

P. A. Fishwick. 2008. *Aesthetic Computing*. MIT Press, Cambridge, MA.

P. Galanter. 2012. Computational aesthetic evaluation: Steps towards machine creativity. In *Proceedings of ACM SIGGRAPH Courses (SIGGRAPH'12)*. ACM Press, New York, NY, Article 14, 162 pages. DOI:http://dx.doi.org/10.1145/2343483.2343497

E. Garces, A. Agarwala, D. Gutierrez, and A. Hertzmann. 2014. A similarity measure for illustration style. *ACM Trans. Graph.* 33, 4, Article 93 (July 2014), 9 pages. DOI:http://dx.doi.org/10.1145/2601097.2601131

X. C. He and N. H. C. Yung. 2004. Curvature scale space corner detector with adaptive threshold and dynamic region of support. In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)*. IEEE, WaShington, DC, 791–794. DOI:http://dx.doi.org/10.1109/ICPR.2004.1334377

P. W. Henderson and J. A. Cote. 1998. Guidelines for selecting or modifying logos. *J. Market.* 62, 2 (Apr. 1998), 14–30. DOI:http://dx.doi.org/10.2307/1252158

F. Hoenig. 2005. Defining computational aesthetics. In *Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging (CAe'05)*. Eurographics Association Aire-la-Ville, Switzerland, 13–18. DOI:http://dx.doi.org/10.2312/COMPAESTH/COMPAESTH05/013-018

Y. Ke, X. Tang, and F. Jing. 2006. The design of high-level features for photo quality assessment. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*. IEEE, Washington, DC, 419–426. DOI:http://dx.doi.org/10.1109/CVPR.2006.303

L. F. Laursen, Y. Koyama, H. T. Chen, E. Garces, D. Gutierrez, R. Harper, and T. Igarashi. 2016. Icon set selection via human computation. In *Pacific Graphics Short Papers (PG'16)*. The Eurographics Association. DOI:http://dx.doi.org/10.2312/pg.20161326

C. Li and T. Chen. 2009. Aesthetic visual quality assessment of paintings. *IEEE J. Select. Top. Signal Process.* 3, 2 (Apr. 2009), 236–252. DOI:http://dx.doi.org/10.1109/JSTSP.2009.2015077

Y.-N. Li, K. Zhang, Y. J. Fu, and D. J. Li. 2014. Rule-based automatic generation of logo designs. *Leonardo* 0, 0 (Nov. 2014), 173–177. DOI:http://dx.doi.org/10.1162/LEON_a_00961

W. H. Liao and P. M. Chen. 2014. Analysis of visual elements in logo design. In *Proceedings of the 12th International Symposium on Smart Graphics (SG'14)*. Springer, Switzerland, 73–85. DOI:http://dx.doi.org/10.1007/978-3-319-11650-1_7

P. Obrador, M. A. Saad, P. Suryanarayan, and N. Oliver. 2012. Towards category-based aesthetic models of photographs. In *Proceedings of the 18th International Conference on Multimedia Modeling (MMM'12)*. Springer, Switzerland, 63–76. DOI:http://dx.doi.org/10.1007/978-3-642-27355-1_9

P. O'Donovan, J. Libeks, A. Agarwala, and A. Hertzmann. 2014. Exploratory font selection using crowdsourced attributes. *ACM Trans. Graph.* 33, 4, Article 92 (July 2014), 9 pages. DOI:http://dx.doi.org/10.1145/2601097.2601110

D. L. Page, A. F. Koschan, S. R. Sukumar, B. Roui-Abidi, and M. A. Abidi. 2003. Shape analysis algorithm based on information theory. In *Proceedings of the International Conference on Image Processing (ICIP'03)*. IEEE, Washington, DC, 229–234. DOI:http://dx.doi.org/10.1109/ICIP.2003.1246940

R. Rauschenberger, J. J.-W. Lin, X. S. Zheng, and C. Lafleur. 2009. Subset search for icons of different spatial frequencies. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting (HFES'09)*. SAGE, Thousand Oaks, CA, 1101–1105.

R. Reber, N. Schwarz, and P. Winkielman. 2004. Processing fluency and aesthetic pleasure: Is beauty in the perceivers processing experience? *Pers. Soc. Psychol. Rev.* 8, 4 (Feb. 2004), 364–382. DOI:http://dx.doi.org/10.1207/s15327957pspr0804_3

K. Reinecke, T. Yeh, L. Miratrix, R. Mardiko, Y. Zhao, J. Liu, and K. Z. Gajos. 2013. Predicting users' first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'13)*. ACM, New York, NY, 2049–2058. DOI:http://dx.doi.org/10.1145/2470654.2481281

B. Saleh, M. Dontcheva, A. Hertzmann, and Z. Liu. 2015. Learning style similarity for searching infographics. In *Proceedings of the 41st Graphics Interface Conference (GI'15)*. Canadian Information Processing Society, Toronto, Canada, 59–64. http://dl.acm.org/citation.cfm?id=2788890.2788902

A. Sartori, V. Yanulevskaya, A. A. Salah, J. Uijlings, E. Bruni, and N. Sebe. 2015. Affective analysis of professional and amateur abstract paintings using statistical analysis and art theory. *ACM Trans. Interact. Intell. Syst.* 5, 2, Article 8 (July 2015), 27 pages. DOI:http://dx.doi.org/10.1145/2768209

C. E. Shannon. 1948. A mathematical theory of communication. *Bell Syst. Techn. J.* 27, 3 (Oct. 1948), 379–423. DOI:http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x

S. Thumfart, R. H. A. H. Jacobs, E. Lughofer, C. Eitzinger, F. W. Cornelissen, W. Groissboeck, and R. Richter. 2011. Modeling human aesthetic perception of visual textures. *ACM Trans. Appl. Percept.* 8, 4, Article 27 (November 2011), 29 pages. DOI:http://dx.doi.org/10.1145/2043603.2043609

C. Wallraven, D. Cunningham, and J. Rigau. 2009. Aesthetic appraisal of art-from eye movements to computers. In *Proceedings of the International Symposium on Computational Aesthetics in Graphics, Visualization and Imaging (CAe'09)*. Eurographics Association, Switzerland, 137–144.

Q. Wang. 2009. Statistical analysis on the enterprise logo color designs of global 500. In *Proceedings of the IEEE 10th International Conference on Computer-Aided Industrial Design&Conceptual Design (CAID&CD'09)*. IEEE, Washington, DC, 277–280. DOI:http://dx.doi.org/10.1109/CAIDCD.2009.5375017

A. W. White. 2011. *The Elements of Graphic Design* (2nd. ed.). Allworth Press, New York, NY.

L. K. Wong and K. L. Low. 2009. Saliency-enhanced image aesthetics class prediction. In *Proceedings of the IEEE International Conference on Image Processing (ICIP'09)*. IEEE, Washington, DC, 997–1000. DOI:http://dx.doi.org/10.1109/ICIP.2009.5413825

X. S. Zheng, I. Chakraborty, J. J.-W. Lin, and R. Rauschenberger. 2009. Correlating low-level lmage statistics with users' rapid aesthetic and affective judgments of web pages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)*. ACM, New York, NY, 1–10. DOI:http://dx.doi.org/10.1145/1518701.1518703