

A Model to Compute Webpage Aesthetics Quality Based on Wireframe Geometry

Ranjan Maity^(✉) and Samit Bhattacharya

Department of Computer Science and Engineering,
Indian Institute of Technology, Guwahati, Guwahati 783019, Assam, India
{ranjan.maity, samit}@iitg.ernet.in

Abstract. Computational models of web page aesthetics prediction are useful for the designers to determine the usability and to improve it. Positional geometry of webpage objects (wireframe) is an important factor for determining the webpage aesthetics as shown in studies. In this paper, we propose a computational model for predicting webpage aesthetics based on the positional geometry. We have considered 13 features of positional geometry that affect aesthetics, as reported in literature. By varying these 13 features, we have designed 52 interfaces' wireframes and rated them by 100 users in a 5 point rating scale. Our 1 dimensional *ANOVA* study on users' rating shows, 9 out of the 13 features are important for webpage aesthetics. Based on these 9 features, we created a computational model for webpage aesthetics prediction. Our computational model works based on *Support Vector Machine (SVM)*. To judge the efficacy of our model, we considered 10 popular webpages' wireframes, and got them rated by 80 users. Experimental results show that our computational model can predict webpage aesthetics with an accuracy of 90%.

Keywords: Computational model · Aesthetics · Features · Empirical study · Classification · ANOVA · Support vector machine

1 Introduction

With the advancement of technology, electronic devices like mobiles, laptops, i-pads etc. have become an essential part in our daily life. As a result, design of interfaces has become a vital issue for the researchers of Human Computer Interaction (HCI) community. Good interface design requires knowledge about people: how they see, understand and think [5]. Another important aspect of good design is how information is visually presented to the users or the *aesthetics* of the interface. According to the Oxford dictionary [17], *aesthetics* is “concerned with beauty and art and the understanding of beautiful things”. It is argued that aesthetically pleasing interfaces increase user efficiency and decrease perceived interface complexity, which in turn helps in increasing usability, productivity and acceptability [15]. It has been also observed [2] that one redesigned graphical window, can save \$20,000 of company during the first year of use. The general rule of thumb says every dollar invested in usability returns \$10 to \$100 [16]. So the current requirement of interface design is the design of aesthetically pleasing interface.

Thus, in general we can say that the interface aesthetics have an important role to play in determining usability. This is true particularly in the web page design. A webpage is composed of three different types of basic elements, or any combination of them, namely: text, image and short video/animation. Other webpage elements like: icon, table, and link etc. can also be approximated by text and image; like icon with image, table and link with text. The shapes of all the webpage elements can be approximated by rectangles. For example, a circular image can be treated as content of a rectangle (*wireframe*).

Since aesthetics is important in webpage design, it is necessary to measure it. Aesthetic measurement of webpage is considered to be *subjective*. Hence, the measure is primarily done through empirical means. A parallel research effort also attempted to develop computational models to evaluate aesthetics of the whole interface [12]. The advantage of computational model is the ability to evaluate interface aesthetics automatically, which in turn makes it possible to automate the design process itself, as demonstrated in [14].

Although a number of works [1, 4, 6–11] report on the contents of a webpage, Ngo et al. [12] found that the impacts of the contents on aesthetics are not important. Rather, the positional geometry of the webpage elements (wireframe) is strongly related with aesthetics. They proposed a computational model for webpage aesthetics based on the 13 positional geometry features of webpage elements. The average of these 13 feature values (termed as *order*) was used to judge the aesthetics of a webpage. Again Lai et al. [7] claimed that the *symmetry* feature had no relation with aesthetics. Finally, Ngo et al. [12] claimed that, their model may not be appropriate for the real webpages. So, we felt that there is a need to review the geometry related features and to develop a computational model for webpage aesthetics.

In this work, we re-examined the 13 best known features of webpage aesthetics, as reported in [12]. We created wireframes of 52 webpages and rated them by 100 participants. ANOVA study on the users' ratings revealed only 9 out of the 13 features are statistically significant for aesthetics measurement. Based on those 9 features, we developed a computational model to predict webpage aesthetics. Our model work based on the *linear kernel* of Support Vector Machine (SVM) [3]. To judge the efficiency of our model, we considered the wireframes of 10 real webpages, and rated them by 80 users. Experimental results show that our model can predict aesthetics with 90% accuracy. The details of the 13 positional geometry features, empirical data collections, the proposed model and analysis are described in this paper.

2 Empirical Study for Feature Identification

In order to find the impact of the 13 independent features associated with aesthetics [12], we performed an empirical study. We created 4 webpages for each feature by systematically varying the feature values. These webpages were rated by 100 participants. One dimensional ANOVA was used to find the impact of each feature, associated with webpage aesthetics.

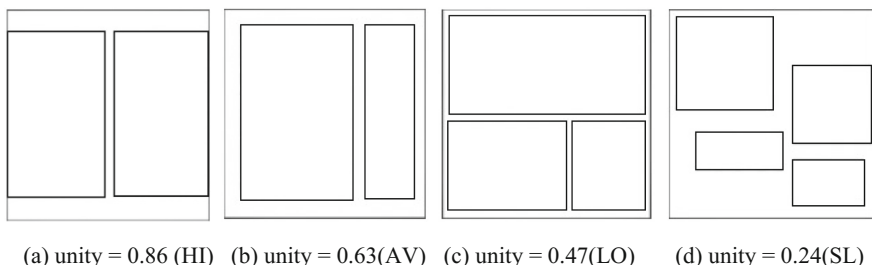
For each feature, we created 4 webpages by varying the feature values systematically in 4 levels – *significant low* (SL), *low* (LO), *average* (AV), *high* (HI). Range for the feature values for each level is shown in Table 1. All these features are independent

Table 1. Four feature class with their range

Feature class	Min value	Max value	Feature class	Min value	Max value
SL	0	0.25	LO	>0.25	0.5
AV	>0.5	0.75	HI	>0.75	1

of each other, as reported in [12]. During the variation of each feature (in 4 different feature classes), we did not observe any significant variations in the other features, which may affect aesthetics.

Altogether, we designed 52 webpage wireframe models using Adobe Photoshop CS6TM. The size of each model was 700×700 pixels. Figure 1 shows set of 4 such models, where the *unity* feature varies. All these 52 models were shown to the participants on PCs having 2.6 GHz AMD Phenom II X3 710, processor running on Windows 8. Each PC had a 23 in. wide viewing angle color display.

**Fig. 1.** Wireframe models of webpages designed by varying the unity feature

One hundred participants took part in our study. Out of them, 20 were school students (average age 16 years), 60 were under and post graduate students (average age 22 years), and rest 20 were teachers (average age 36 years). All of them had normal or corrected-to-normal vision and none of them was colour blind (self reported). All of them were regular computer users. However, none was familiar with screen design concepts.

We created webpages' model of the 52 webpages without considering the contents of the webpage elements. All the 52 models were rated by 100 users in a five point rating scale (1–5); *five* denoted aesthetically pleasing webpage, and *one* denoted the aesthetically least pleasing webpage. A browser-based viewer was created for the users, with facilities to view previous/next sample and to rate a webpage model. After viewing each webpage model, participants rated it according to its *aesthetic appeal*. Each participant rated the 52 models assigned to him/her in two sessions (26 each) in a day. They were allowed to take breaks in each session. These measures were taken to avoid discomfort to the participants that might have arose due to the large number of webpages to be rated. To avoid the learning effect, we randomly varied the sequence of the webpage models shown to the users. Before data collection, we performed a small

training sessions for the participants. In these sessions, participants were familiarized with the 5 point scale and the web interface by which they had to rate the webpage models.

Results and Analysis

We computed the feature values of the 52 webpages using the analytical expressions proposed by Ngo's work [12].

Based on the empirical results, we performed 13 independent 1 dimensional ANOVA (one for each feature) by using the ANOVA1 command of MATLAB 2014. We observe that the p values are higher ($p > 0.05$) for the features - *density* (0.11), *economy* (.108), *rhythm* (.174) and *simplicity* (.0536). This implies that the variations of these feature values were not having statistically significant impact on the webpage aesthetics. On the contrary, the remaining 9 features – *balance*, *cohesion*, *equilibrium*, *homogeneity*, *proportion*, *regularity*, *sequence*, *symmetry* and *unity* were found to be statistically significant for the webpage aesthetics. So, we consider these 9 geometry features for webpage aesthetics. Accordingly, we propose a computational model based on these 9 features, as discussed below.

3 Proposed Computational Model

Our model works based on the *linear kernel* of *Support Vector Machine* (SVM) [3]. SVM is popularly used for solving the binary classification problems. In the following section, we discuss about the training procedure of our model.

3.1 Model Training

For the model training, we considered a subset of data (9 out of 13 features). It may be noted that we have 5200 ($100 \times 4 \times 13$) data points which are the ratings of 100 users on the 52 webpages; 4 webpages for each of the 13 features. Out of these 13 features, we considered 9 features for the training of our model. So, for these 9 features we have 36 webpages (4 for each feature) and their corresponding 100 users' ratings. Altogether, we have 3600 ($9 \times 4 \times 100$) training data points. In our model we consider 9 SVMs to predict the 9 features independently. Each of these 9 SVMs was trained by the 400 data points (100×4), which are the ratings of 100 users for the 4 different webpages of a particular feature. As, SVM works on labeled data, we converted all the 3600 unlabeled data to labeled training data by using the following logic.

Data Labeling

Input: 36 feature values and their users' rating (100)
Output: feature class (good or bad)

```

if (feature value >= 0.5 and users rating >= 3)
    feature class = good (labeled by +1)
else
    feature class = bad (labeled by -1)
end if

```

A particular feature of a webpage was labeled as good (+1), only when an user gave a rating more than 2, in a 5 point scale, as well as the feature value was greater than or equals to 0.5 (computed by the Ngo’s analytical expression [12]) in 0 to 1 scale. This was done to label almost half of the scale for aesthetically pleasing (good) feature and the rest was for aesthetically unpleasing feature. Based on these labeled data, we trained our model using the *SVMTRAIN* function of *MATLAB* 2014.

3.2 Empirical Study for Model Validation

The main objective of our model is to predict aesthetics for real webpages. For this purpose, we performed another empirical study on 10 real webpages. The details of the empirical study, along with the performance of our model are discussed below.

For model validation, we considered wireframes of 10 real webpages [18–27] (*home* pages of 10 websites). These webpages represent some popular domains, like – education, banking, e-commerce, social networking, entertainment, news and corporate sector. For all these webpages, we created the webpage models (constructed without considering the content of the webpage elements) by using *Adobe Photoshop CS6*TM. Figure 2(b) shows the model of the *soundcloud* (Fig. 2(a)) page. These models were rated by the users in PCs, having 2.6 GHz AMD Phenom II X3 710 processor, running on *Windows* 8. Each PC had a 23 in. wide viewing angle colour display.

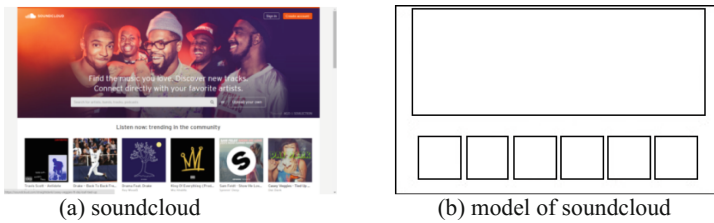


Fig. 2. Webpages and their model

All the 10 webpage models were rated by 80 new users. Out of them, 40 were male and the rest were female. Fifty participants were undergraduate students (average age 21 years), 20 users were postgraduate students (average age 27 years) and rest 10 were the faculty members (average age 40 years). All the participants were regular computer users but none of them had any knowledge about the website design principles. All the users had normal or corrected-to-normal vision and none of them was colour blind (self reported).

Each participant rated the 10 webpage models using the same browser-based viewer used in our previous study. After viewing each webpage model, they rated it according to its *aesthetic appeal* in the same 5 point rating scale, used in our previous study; 5 denoted aesthetically pleasing webpage, and 1 denoted aesthetically least pleasing webpage. To avoid the learning effect, we randomly varied the sequence of the webpage models shown to the users. A participant rated the 10 webpages assigned to her/him in one session in a day. Before data collection, we performed a small training

sessions for the participants. In these sessions, participants were familiarized with the 5 point scale and the web interfaces by which they had to rate the webpages.

3.3 Results and Discussion

The 13 feature values and their average of the 10 webpage models were computed with the help of Ngo's [12] formulas. Although we selected the 9 features, but computation of these 13 feature values will help us to prove the efficiency of our model. Table 2 shows the *order* (average values of all the 13 features) of the 10 webpages' wireframes. Using these wireframes, we performed another empirical study. Results of the empirical study are shown in Table 2. The mode column of Table 2 denotes the rating, given by most of the users for a particular webpage. Based on the mode value, we classified the webpages in two classes – good or bad. For the binary classification we used the following logic.

```

if mode >= 3
    webpage class=good(aesthetically pleasing(+1))
else
    webpage class=bad(aesthetically unpleasing(-1))

```

We independently trained our *SVMs* to predict the feature class (good or bad) of the 9 features (for 10 webpages), reported in Table 2. Feature class prediction was done using the *SVMPREDICT* function of *MATLAB* 2014 version. Finally, for predicting the aesthetics of the whole webpage we used the following algorithm.

Prediction Algorithm for a webpage

```

Input: 9 feature class values (good or bad), count = 0
Output: webpage class (good or bad)
for i = 1 to 9
    if (feature class == good)
        count = count+1
    endif
    i=i+1
end for
if (count>=5)
    webpage class = good (labeled by +1)
else
    webpage class = bad (labeled by -1)
endif

```

The prediction algorithm shows, if most of the predicted features (5 out of 9) are aesthetically pleasing (good) then our algorithm (as mentioned above) predicts the webpage as good (aesthetically pleasing), otherwise it is treated as bad (aesthetically unpleasing).

We compared the predicted result of our model with the results obtained from empirical study. Using our model, we predicted the feature types of the 10 webpages as shown in Table 2. Then based on our webpage prediction algorithm we predicted the webpages as aesthetically pleasing (+1) or not (-1). The second rightmost column of the Table 2 shows the result of the webpage type (aesthetically pleasing or not) by the users' rating.

Table 2. Empirical study result vs. predicted result

Website	Label predicted by our developed SVM									Model predicted webpage label	Mode of users' ratings	Users' choice webpage label	Order value based on Ngo [18]
	Balance	Cohesion	Equilibrium	Homogeneity	Proportion	Regularity	Sequence	Symmetry	Unity				
CIT, Kokrajhar	-1	-1	+1	-1	+1	+1	+1	-1	-1	-1	3	+1	0.4706
Sound Cloud	+1	+1	+1	-1	+1	+1	+1	-1	+1	+1	4	+1	0.5216
SBI	-1	-1	+1	-1	-1	-1	-1	-1	-1	-1	2	-1	0.4334
Flipkart	-1	-1	+1	-1	-1	-1	-1	-1	-1	-1	2	-1	0.4775
Facebook	-1	-1	+1	-1	+1	+1	+1	+1	-1	+1	3	+1	0.4532
Amazon	+1	+1	+1	-1	+1	+1	+1	-1	+1	+1	4	+1	0.5133
NewsLive	-1	-1	+1	-1	-1	-1	-1	-1	-1	-1	2	-1	0.4520
Youtube	-1	-1	+1	-1	-1	-1	-1	-1	-1	-1	2	-1	0.4406
IIT Guwahati	+1	+1	+1	-1	+1	+1	+1	-1	-1	+1	4	+1	0.5003
TCS	+1	+1	+1	-1	+1	+1	+1	-1	-1	+1	4	+1	0.4581

Out of 10 webpages, our model accurately predicted 9 webpages. Thus, our model predicted webpage aesthetics with an accuracy of 90% (9 out of 10).

Ngo et al. [12] claimed that the *order* value may be a measure for aesthetics computation. However, in our study we observe that the *order* may not be relevant for real webpages. The *order* value of the 10 real webpages lied in the range from 0.43 to 0.52 as shown in Table 2, which is sorted based on the *order* value. It may be noted that, the order value of the *facebook* is 0.45 was treated by users as aesthetically pleasing webpage. In a contrary, the higher *order* value (than that of *Facebook*) of *Flipkart* 0.47 was treated as aesthetically displeasing by the users. Again, the lower *order* value (than that of *Flipkart*) of *TCS* and *CIT, Kokrajhar* are 0.4532 and 0.4581 were marked as aesthetically pleasing webpages by the users. Even we can observe that the difference of the *order* value among the webpages *NewsLive* and *Facebook* is only 0.1%. But, still users find *facebook* as aesthetically pleasing, while the *NewsLive* as aesthetically displeasing. Similarly, we observe the difference in the *order* value is only 0.7% among *CIT Kokrajhar* and *Flipkart*. But *CIT Kokrajhar* was treated as aesthetically pleasing, while the *Flipkart* was treated as aesthetically displeasing.

Based on the above observations, we can claim that *order* is not a suitable metric for aesthetics computation of real webpages. In contrast, our model works based on *SVM*; which has the capability for solving binary classification problems with high accuracy.

SVM maps data into a higher dimensional input space and creates an optimal separating hyper plane in the higher dimensional space. As a result, two classes (good or bad) are created across the separating hyper plane. Then for a particular input, *SVM* predicts the corresponding class. We used this notion to predict the feature types of the 10 webpages. However, selection of *SVM* kernel is a tricky task. The general convention is to use the linear kernel first, as they are easier and faster than that of the others kernels, like *polynomial*, *RBF* and *Sigmoid* kernels. If the prediction result is satisfactory then *linear* kernel is the best option; otherwise, other kernels have to be used. Using this convention, we used the *linear* kernel of *SVM* in our model. We

observed an accuracy of 90% in aesthetics prediction, which is good enough. So, we refrain to consider the other kernels. However, the performances of the other kernels in this context may be explored. In our model, we trained each *SVM* by 400 data. However, larger training data may improve the performance of our model.

Our model can predict webpage aesthetics in terms of two classes – aesthetically pleasing (good) or unpleasing (bad). It is the simplest type of classification and has the less chance of misclassifying than that of multiclass classifiers. Again, multiclass classifiers are more complex and consume more time for classification, than that of binary classifiers. So, for the devices, where time and computational power are the judgement factors, like – mobiles, PDAs etc., binary classifier is the best choice. In our study we collected the users’ rating in a 5 point rating scale and predicted webpage aesthetics in two classes (good or bad). However, the performance of a model which can predict aesthetics in same scale (1–5) may be explored for further analysis.

Our computational model can help the designers to improve the design. For a particular webpage if our model predicts it as aesthetically unpleasing, then there is likely to be some problem with the design, which can reduce aesthetics and consequently, reduce usability of the design. Hence, the designer should take some corrective measures to improve the design. A designer can correct his/her design by considering the predicted *feature_class* values produced by our model. We can integrate our model with the webpage design guidelines or apply genetic algorithm based approach as reported in [14] to redesign webpage geometry. This redesign can improve aesthetic appeal; which in turn can increase usability.

In this work we have only considered the size and geometrical positions of the webpage elements for aesthetics. We have not considered the content of the webpage elements. But in some studies we observed that contents of webpage elements have impact on aesthetics. Our model may be combined with the models of short animations [14], text [10] and image model [9] to develop a complete computational model for evaluating aesthetic quality of a webpage. In our work, we consider the 4 variations of each feature. However, making more variations in each feature value may help us to develop a better model.

4 Conclusion

In this work, we reassessed the best known features for webpage aesthetics. We performed an empirical study and found that 9 features are important for aesthetics measurement. By considering these 9 features, we developed a computational model for aesthetics prediction. Our model works based on the *linear kernel* of Support Vector Machine. To judge the efficiency of our model, we performed another empirical study on real webpages, and found that our model can predict webpage aesthetics with a high accuracy of 90%. In future, we plan to refine and extend our model by more empirical data collection for model training and testing. We also try to combine our model with other predictive models for text, image and short animation aesthetics and develop a fully automated design environment. Investigation of more robust model based on other *kernels* of *SVMs* may also be explored. In our model we predicted

aesthetics as good or bad. However, predicting aesthetics in the 5 point rating scale may also be an interesting topic. For each feature we considered only 4 variations. More variations in the feature values may help to build a more robust model.

Acknowledgement. We are thankful to all the participants who volunteered for the empirical studies.

References

1. Bansal, D., Bhattacharya, S.: Semi-supervised learning based aesthetic classifier for short animations embedded in web pages. In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) INTERACT 2013. LNCS, vol. 8117, pp. 728–745. Springer, Heidelberg (2013). doi:[10.1007/978-3-642-40483-2_51](https://doi.org/10.1007/978-3-642-40483-2_51)
2. Cope, M.E., Uliano, K.C.: Cost-justifying usability engineering: a real world example. In: Proceedings of the Human Factors Society 39th Annual Meeting, vol. 39, no. 4, pp. 263–276. Human Factors Society, Santa Monica (1995)
3. Cortes, C., Vapnik, V.: Support-vector networks. *Mach. Learn.* **20**(3), 273–297 (1995)
4. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Studying aesthetics in photographic images using a computational approach. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3953, pp. 288–301. Springer, Heidelberg (2006). doi:[10.1007/11744078_23](https://doi.org/10.1007/11744078_23)
5. Galitz, W.O.: *The Essential Guide to user Interface Design: An Introduction to GUI Design Principles and Techniques*. Wiley, New York (1997)
6. Hill, A.: Readability of screen displays with various foreground/background color combinations, font styles, and font types. In: Proceedings of 11th National Conference on Undergraduate Research, vol. 2, pp. 742–746 (1997)
7. Lai, C.Y., Chen, P.H., Shih, S.W., Liu, Y., Hong, J.S.: Computational models and experimental investigations of effects of balance and symmetry on the aesthetics of text-overlaid images. *Int. J. Hum. Comput Stud.* **68**, 41–56 (2010)
8. Legge, G.E., Bigelow, C.A.: Does print size matter for reading? A review of findings from vision science and typography. *J. Vis.* **11**(5), 8 (2011)
9. Maity, R., Uttav, A., Verma, G., Bhattachariya, S.: A non-linear regression model to predict aesthetics ratings of on-screen images. In: OZCHI 2015, Melbourne, Australia, pp. 44–52 (2015)
10. Maity, R., Madrosia, A., Bhattachariya, S.: A computational model to predict aesthetics quality of text elements of GUI. In: IHCI 2015, Allahabad, India, pp. 152–159 (2015)
11. Miniukovich, A., Angeli, D. Computation of interface aesthetics. In: CHI 2015, Seoul, Republic of Korea, pp. 1163–1172 (2015)
12. Ngo, D.C.L., Teo, L.S., Byrne, J.G.: Modeling interface aesthetics. *Inf. Sci.* **152**, 25–46 (2003)
13. Shyam, D., Bhattacharya, S.: A model to evaluate aesthetics of short videos. In: APCHI, Matsue, Japan, pp. 315–324 (2012)
14. Singh, N., Bhattacharya, S.: A GA-based approach to improve web page aesthetics. In: Proceedings of 1st international conference on Intelligent Interactive Technologies and Multimedia (IITM 2010), pp. 29–32. IIT Allahabad, India (2010)
15. Tractinsky, N.: Aesthetics and apparent usability: empirically assessing cultural and methodological issues. In: CHI 1997, New York (1997)
16. International Business Machines (IBM). Cost Justifying Ease of Use (2001)
17. http://www.oxfordlearnersdictionaries.com/definition/english/aesthetic_1

18. [CIT Kokrajhar] cit.ac.in
19. [IIT Guwahati] iitg.ac.in
20. [Facebook] [facebook.com](https://www.facebook.com)
21. [Sound cloud] [soundcloud.com](https://www.soundcloud.com)
22. [Youtube] [youtube.com](https://www.youtube.com)
23. [Flipkart] [flipkart.com](https://www.flipkart.com)
24. [State Bank of India] [onlinesbi.com](https://www.onlinesbi.com)
25. [Amazon] [amazon.com](https://www.amazon.com)
26. [Newslive] [newslivetv.org](https://www.newslivetv.org)
27. [TCS] [tcs.com](https://www.tcs.com)