Effects of complexity on aesthetic preference in webpage design

Dongmei Gao
dongmei.gao@aalto.fi
Aalto University
Espoo, Finland

Abstract

Designers seek to create designs that elicit aesthetics responses from users. Complexity is one of important parameters. HCI researchers have investigated the complexity-aesthetics relationship in recent years but showed different results, including an inverse linear correlation, an inverted u-shaped curve and no relationship at all. However, the definition of complexity is still unclear and there are other confounders that have an unknown effect on the complexity-aesthetics relationship. Here I re-analyzed complexity and aesthetics ratings collected in past studies. The results showed common features of complexity are more colorful, less symmetrical and have more sections. The second study showed that novelty and craftsmanship (webpages designed by professional or non-professional designers) had different effects on the relationship. After factoring them out, the complexity-aesthetics relationship became no relationship or negative linear . The results demonstrate how complex webpages look like and how the complexity of aesthetics influences the perceived aesthetics by humans. It can provide designers with guidance on web design, how to balance complexity and aesthetics.

CCS Concepts: • **Computer systems organization** → **Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability.

Keywords: datasets, neural networks, gaze detection, text tagging

ACM Reference Format:
Dongmei Gao. 2021. Effects of complexity on aesthetic preference in webpage design. In *Woodstock '21: ACM Symposium on Neural Gaze Detection, May 14–20, 2021, Woodstock, NY*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/1122445.1122456

Woodstock '21, May 14–20, 2021, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

1 Introduction

When talking about webpage design, a websites' first impression is known to be a crucial moment for capturing the users' interest. Visual complexity is considered an important factor for aesthetic perception of websites in the context of first impression and low complexity usually creates a better first impression[3]. In the context of human-computer interaction or interactive product design, the complexity together with other attributes will affect the emotional response of users. Complex interface design is considered to be more likely to cause emotional arousal and valence than simple interface design[4].

As a result, complexity plays an important role in webpage design. There is a lot of related research about the relationship between complexity and aesthetics in webpage design, but until now no exact definition is concluded. One possible reason is that both complexity and aesthetics are subjective feelings without an exact definition, therefore, different data and methods will lead to different results.

In this paper, I analyzed the existing datasets from a previous complexity and aesthetics research. It includes scores of webpage complexity, aesthetics, novelty and craftsmanship. I used several common metrics to test the 32 web pages with the highest and lowest complexity scores. Those metrics will make the definition of complexity more clear and understandable. The second study was conducted to study the aesthetics-complexity relationship. By building regression models and studying the correlation between complexity, aesthetics, novelty and craftsmanship, I got how novelty and craftsmanship affected the aesthetics-complexity relationship. The results recommend a design principle for designers to take complexity, novelty and craftsmanship into account when they are designing webpages. Designers are able to know how to balance those metrics to make the webpages more aesthetic. There are various possibilities of a combination of complexity and aesthetics in the design field.

2 Related Work

Berlyne's psychobiological theory posited the inverted-U model to explain the relationship between aesthetics impression and stimulus properties, such as complexity, familiarity and novelty. This is one of the well-known theories which attracts a lot of researchers to prove it in different fields. Two opposing neurological systems work together to affect the

model. Take complexity as an example, the reward system responds to initial increases of the degree of complexity and leads the aesthetics impression increasing. When the complexity reaches a certain level, another system – aversion system, becomes activated but drives the aesthetics impression in the opposite direction if the complexity keeps rising.

However, Miniukovich et al.[1] argued the completeness of content would affect the relationship. If some content in webpage was missing, the decline in aesthetics is not only related to complexity. So they conducted a study to prove if accounting for technical condition, the relationship becomes largely linear instead of inverted-U correlation. There are still other factors that may eliminate the u-shaped or linear tendency of the complexity-aesthetics relationship.

According to existing research, some focused on proving the Berlyne's theory but the definition of complexity is ambiguous. Although human beings' judgment on complexity levels is considered to be congruous[2], computer program can't understand the human emotions. When considering using computer program to automatically judge complexity degree of webpages, a clear definition of complexity is necessary. One possible summary of visual complexity features is including information about the organization, statistics, distribution of one image, etc.

3 PRE-STUDY

3.1 Datasets

The dataset I got from previous research contains 6 text files and 1506 screenshots of webpages. One text file include demographic information of participants, with different participants differentiated using IDs. These IDs correspond to the column names in the other 5 text files, which contain participants' evaluations of webpage visual complexity, aesthetics, technical condition, novelty and design craftsmanship. Each participant evaluated only 1 of these 5 webpage qualities and only for 100 out of 1506 webpages - the missing scores are put as NA.

The 1506 screenshots are a combination of 6 previously published and 1 new dataseet. The past datasets are (screenshots names begin accordingly):

IJHCS_12: Tuch, A. N., Presslaber, E. E., StöCklin, M., Opwis, K., & Bargas-Avila, J. A. (2012). The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. International journal of human-computer studies, 70(11), 794-811.

CHI_13: Reinecke, K., Yeh, T., Miratrix, L., Mardiko, R., Zhao, Y., Liu, J., & Gajos, K. Z. (2013, April). 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 (pp. 2049-2058). ACM.

Figure 1. Webpages with the highest(left) and the lowest(right) complexity score

AVI_14: Miniukovich, A., & De Angeli, A. (2014, May). Quantification of interface visual complexity. In Proceedings of the 2014 international working conference on advanced visual interfaces (pp. 153-160). ACM.

CHI_15: Miniukovich, A., & De Angeli, A. (2015, April). Computation of interface aesthetics. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 1163-1172). ACM.

ICWE_19: Boychuk, E., & Bakaev, M. (2019, June). Entropy and Compression Based Analysis of Web User Interfaces. In International Conference on Web Engineering (pp. 253-261). Springer, Cham.

CHI_20 is a prefix for the new, previously not used/published webpages

All webpages are identified using artificial, numerical IDs, and are not explicitly linked to the websites they were collected from.

4 Approach

According to the complexity scores, I got the webpages with an average score higher than or equal to 6. There are 16 webpages shown in Table 1. And as a comparison, Table 1 showed 16 webpages with the lowest complexity score. Figure 1 showed two webpages with the highest score and the lowest score separately. I used the website https://interfacemetrics.aalto.fi/ to compute each webpage's metrics and models that predict how users perceive, search, and experience the design. This website includes four parts of metrics: Color Perception, Perceptual Fluency, Visual Guidance and Accessibility.

I mainly focused on four metrics which can represent the complexity difference well to a certain extend:

Unique RGB colors: The number of unique colours in RGB spectrum is an indication of colour variance. Colours that occur more than a threshold value are counted.

Edge Congestion: Edge congestion indicates the ease with which main edges can be perceived. A crowded image is hard to follow. The edge congestion indicator is important for complex interfaces and graph visualizations.

stimulus_id	average_score
ICWE_19_361.png	6.83
CHI_13_138.png	6.81
CHI_13_252.png	6.42
CHI_20_216.png	6.38
CHI_20_166.png	6.29
CHI_13_56.png	6.27
CHI_13_194.png	6.22
IJHCS_12_art255.png	6.2
CHI_20_238.png	6.06
IJHCS_12_bus37.png	6.05
CHI_13_157.png	6.00
CHI_20_26.png	6.00
CHI_13_309.png	6.00
CHI_15_25.png	6.00
ICWE_19_307.png	6.00
ICWE_19_376.png	6.00

Table 1. Top 16 webpages with the highest complexity score

stimulus_id	average_score
IJHCS_12_art234.png	1.61
ICWE_19_319.png	1.63
AVI_14_31.png	1.64
IJHCS_12_art93.png	1.65
ICWE_19_11.png	1.67
CHI_20_161.png	1.67
CHI_13_50.png	1.67
AVI_14_14.png	1.67
IJHCS_12_bus157.png	1.69
AVI_14_119.png	1.75
CHI_20_78.png	1.75
IJHCS_12_art14.png	1.78
IJHCS_12_art184.png	1.78
AVI_14_110.png	1.78
ICWE_19_166.png	1.79
IJHCS_12_bus218.png	1.80

Table 2. 16 webpages with the lowest complexity scoree

Pixel Symmetry: It indicates perceived symmetricity across an axis. It is associated with the Gestalt principle of symmetry. This metric considers the whole image and finds an axis for maximum symmetry.

Quadtree Decomposition - Number of Leaves: Quadtree Decomposition indicates visual complexity of a scene. It recursively breaks down the image into regions based on entropy in colour and luminance channels. Number of leaves is the total amount of leaves at the end of the recursion. The higher this number, the higher the complexity.

Next, I replicated Miniukovich's computation to study the relationship between complexity and aesthetics. Regression models were built to test the existence, shape and strength of the correlation. The regression model with dataset ID as a moderator, aesthetics as output, complexity and complexity 2^{nd} order polynomial as predictors. To factor out the variance of novelty and craftsmanship from aesthetics and estimate complexity-aesthetics relationship more precisely, novelty and craftsmanship were used as a control variable separately in two models regressing aesthetics on complexity linear and quadratic terms.

5 Data Analysis

5.1 Common Features of Complexity

After uploading all 32 webpages, including high complexity and low complexity, I got the result shown in Table 3 and Table 4. From Table 3, the average number of unique colors is 6322 (fair), more than that of less complex webpages (3982, not colorful). Complex webpages usually have more colors. Human are easily distracted by various colors. For the edge congestion metric, complex webpages have an average rate of 0.34 which is not much different from 0.32 for simple web pages. However, there is a big gap between the two in terms of pixel symmetry metric, 0.57 and 0.99 separately. The symmetrical structure of simple webpages can make the content more concentrated and make it easier for readers to perceive the key ideas on the page. As for the number of leaves which is one of metrics of quadtree decomposition, the score of complex web pages (2603) is more than twice that of simple web pages (1240). Complex web pages can be divided into more smaller modules.

As a result, comparing two webpages, if one of them uses rich colors, does not have a prominent symmetrical structure, and has more sub-modules, then it can be considered more complex.

5.2 Reproduce of Complexity-Aesthetics Relationship

To verify the previous results of the existence, shape and strength of the relationship between complexity (both linear and quadratic terms) and aesthetics, two linear models with dataset ID as a moderator were fitted. Because datasets are from different papers and collected in different situations, they were studied separately as well in this paper. I used the SPSS software tool and Python programming to do the regression.

Pearson's and Partial correlation coefficient are two main methods that I used for the analysis. They measure the statistical relationship, or association, between two variables. According to the partial correlation coefficient, the degree of influence of the complexity variable on the aesthetics variable can be judged. From the Table 5, the models suggested that both negative linear and quadratic (inverted U-shape) components of aesthetics-complexity relationship were present, though the strength of the components and their statistical

stimulus_id	average_score	number of unique colors	edge congestion	pixel symmetry	number of leaves
ICWE_19_361.png	6.83	3790	0.34	0.76	1774
CHI_13_138.png	6.81	10020	0.36	0.57	3853
CHI_13_252.png	6.42	4168	0.29	0.53	1615
CHI_20_216.png	6.38	3609	0.26	0.54	2395
CHI_20_166.png	6.29	6734	0.36	0.46	2743
CHI_13_56.png	6.27	1756	0.43	0.57	1960
CHI_13_194.png	6.22	8775	0.33	0.46	3226
IJHCS_12_art255.png	6.2	4822	0.41	0.60	2743
CHI_20_238.png	6.06	6495	0.31	0.31	3481
IJHCS_12_bus37.png	6.05	9473	0.38	0.5	2200
CHI_13_157.png	6.00	18356	0.33	0.59	3697
CHI_20_26.png	6.00	7472	0.29	0.5	3184
CHI_13_309.png	6.00	4101	0.51	0.57	3079
CHI_15_25.png	6.00	3434	0.34	0.61	2101
ICWE_19_307.png	6.00	3515	0.25	0.7	1510
ICWE_19_376.png	6.00	4631	0.24	0.82	2086

Table 3. Top 16 webpages with the highest average complexity score

stimulus_id	average_score	number of unique colors	edge congestion	pixel symmetry	number of leaves
IJHCS_12_art234.png	1.61	732	0.38	0.70	334
ICWE_19_319.png	1.63	3768	0.30	1.05	1939
AVI_14_31.png	1.64	2670	0.16	0.18	2224
IJHCS_12_art93.png	1.65	9117	0.31	0.42	1327
ICWE_19_11.png	1.67	9091	0.27	0.83	2794
CHI_20_161.png	1.67	3169	0.35	0.46	784
CHI_13_50.png	1.67	1176	0.21	1.85	376
AVI_14_14.png	1.67	2117	0.51	1.6	1822
IJHCS_12_bus157.png	1.69	533	0.34	1.10	253
AVI_14_119.png	1.75	499	0.12	0.36	220
CHI_20_78.png	1.75	8471	0.07	0.48	2173
IJHCS_12_art14.png	1.78	8933	0.44	1.35	1480
IJHCS_12_art184.png	1.78	1792	0.40	2.04	589
AVI_14_110.png	1.78	368	0.37	0.32	136
ICWE_19_166.png	1.79	9013	0.31	0.61	2914
IJHCS_12_bus218.png	1.80	2325	0.44	2.42	475

Table 4. 16 webpages with the lowest average complexity score

significance varied depending on the dataset. According to the significance(p) about the strength, we can get similar results as past researches. IJHCS_12 and CHI_15 showed a linear relationship. CHI_13, AVI_14 and ICWE_19 showed a linear relationship with a U-shaped component. CHI_20 showed no relationship because both linear and quadratic components are weak. Figure 2 show the regression lines of them.

5.3 Confounding Factors

To further disambiguate the complexity-aesthetics relationship, three factors – novelty, craftsmanship, and technical condition had been tested. Within the datasets, there were scores of these factors collected from participants. First of all, using the same methods, I reproduced the analysis of the correlation between the three factors, and their correlation with complexity and aesthetics. Same as previous research, only technical condition tended to correlate negatively with both aesthetics and complexity, while novelty correlated negatively with aesthetics and positively with complexity. Craftsmanship correlate positively with aesthetics but didn't have a obvious correlation with complexity (only show weak negative correlation in ICWE_19 and positive correlation in CHI_13).

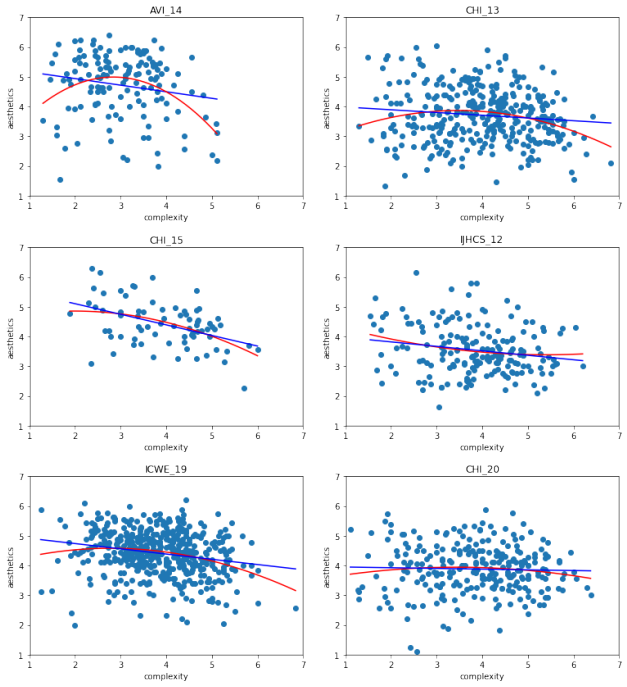

Figure 2. Regression lines show relationships between aesthetics and complexity, blue solid lines are regression lines, red curves are regression lines when $complexity^2$ is added in models

Dataset	Complexity,β(p)	Complexity^2,β(p)
AVI_14	-.18 (.038)	.32 (<.001)
CHI_13	-.11 (.035)	-.17 (<.001)
CHI_15	-.45 (<.001)	-.13 (.263)
IJHCS_12	-.18 (.012)	.07 (.338)
ICWE_19	-.21 (<.001)	-.13 (.004)
CHI_20	.03 (.6)	.08 (.217)

Table 5. A series of regression models, with aesthetics as a dependent variable, and complexity and complexity^2 as predictors; a model was fitted for each dataset, which coefficients as a linear (Complexity) or quadratic (Complexity^2) component of aesthetics-complexity relation, p as the significance of the relation

Next I mainly focused on two of them - novelty and craftsmanship (technical condition was already studied in previous research). Although from the correlation, it seemed they wouldn't affect the relationship between complexity and aesthetics, I wanted to see what would happen after factoring them out.

Novelty. To factor out the variance of novelty from aesthetics and estimate complexity-aesthetics relationship more precisely, novelty was used as a control variable in two models regressing aesthetics on complexity linear and quadratic terms, Table 6. Compared to Table 5, the both linear and the u-shaped component of the relationship became insignificant for all of them but CHI_20. Combining the correlation between novelty and complexity, aesthetics respectively, it can be considered that novelty is an important part of determining complexity. The newer the webpage, the more complex it was. They determined the aesthetics together.

Dataset	Complexity,β(p)	Complexity^2,β(p)
AVI_14	.11 (.21)	-.15 (.07)
CHI_13	.01 (.92)	-.14 (.011)
CHI_15	-.13 (.26)	-.11 (.33)
IJHCS_12	-.01 (.9)	.1 (.18)
ICWE_19	-.02 (.65)	-.13 (.005)
CHI_20	.16 (.009)	.04 (.54)

Table 6. Per-dataset regression models, with residual aesthetics (after factoring out novelty scores) a dependent variance and complexity and complexity^2 as independent variables;

Dataset	Complexity,β(p)	Complexity^2,β(p)
AVI_14	-.23 (.006)	-.21 (.015)
CHI_13	-.26 (<.001)	-.07 (.21)
CHI_15	-.4 (<.001)	-.06 (.62)
IJHCS_12	-.32 (<.001)	.18 (.015)
ICWE_19	-.17 (<.001)	.06 (.16)
CHI_20	-.06 (.32)	.02 (.8)

Table 7. Per-dataset regression models, with residual aesthetics (after factoring out craftsmanship scores) a dependent variance and complexity and complexity^2 as independent variables;

Craftsmanship. As for craftsmanship (webpages designed by professional/non-professional designers), the result varied. After factoring out the craftsmanship factor, the significance of linear component became stronger while that of u-shaped component became weaker. Table 7. According to the results before (the craftsmanship was positively correlated with aesthetics but no relationship with complexity), craftsmanship mainly affect the aesthetics factor. After factoring out craftsmanship, the linear component of complexity-aesthetics relationship became strengthened.

Figure 3 showed regression models of AVI_14 and CHI_13 datasets after factoring out novelty or craftsmanship. Although the negative linear and inverted u-shaped component of the relationship between complexity and aesthetics were present, they were not clear within the effective range (score between 1 to 7).

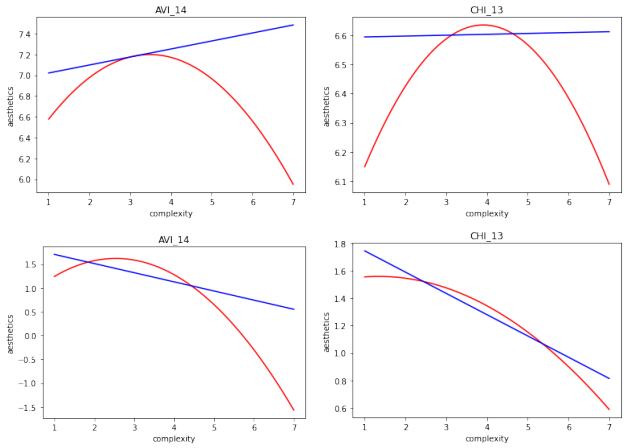

Figure 3. Regression lines of AVI_14 and CHI_13 show relationship between aesthetics and complexity after factoring out novelty factor (top row) or craftsmanship factor (bottom row). Blue solid lines are regression lines, red curves are regression lines when $complexity^2$ is added in models

6 Discussion

This study aimed at determining the common features of complex webpages and studied how novelty and craftsmanship factors affect the relationship between complexity and aesthetics of webpages. Complex webpages are usually more colorful but without a prominent symmetrical structure. They also contains many sections. There are no detailed number of colors or sections because they are more valuable when comparing between webpages. In addition, not individual metrics determine the complexity of a webpage, when I observed the webpages, I found a complex webpage may be less colorful but contain many sections. Multiple metrics affect the complexity of web pages at the same time. Human perceive the complexity from multiple dimensions, so it's difficult to define the complexity in one or two metrics. In future research, when taking complexity into account, it's better to select specific webpages with only one metric varying and other metrics controlled. For example, for 100 webpages, all of them have the same number of colors and similar structures, but the number of sections are changed.

The results of past studies were largely replicated, with both a negative correlation and inverted u-shaped curve supported by different dataset. Novelty has positive correlation with complexity and negative correlation with aesthetics. After controlling it, the significance of linear and u-shape components the complexity-aesthetics relationship became weak. The linear correlation even became positive rather than negative. Novelty has a greater impact on complexity. When we design webpages, if we want it to be more novel, it is also risky to make it more complex, therefore making it less aesthetics. According to craftsmanship, it has positive correlation with aesthetics and no consistent correlation with complexity. After factoring out it, the linear component strengthened while the u-shaped component became weak. The complexity of webpages can not be determined by whether their designers are professional or not. But professional designers have more possibilities to make webpages more aesthetic.

As for Berlyne's theory which claims the inverted-U relationship between complexity and aesthetics, it provided a theoretical basis for many studies but not many studies matched it perfectly. The reason is that this theory required to sample stimuli from the entire simple-complex continuum to observe an inverted U-shaped curve; otherwise, it would find a positive (largely simple stimuli, left side of Berlyne's hypothesized u-shape) or negative (largely complex stimuli, right side of the u-shape) correlation. This may be the reason for the different results obtained in the past research. Another explanation is that for some webpages, participants might have different judgment basis according to their experience. For example, if participants were working in medical industry, they might think related webpages were simpler than others because they were familiar with them. As a result, Berlyne's theory needs a standard sampling which is hard to get in reality.

7 Conclusion

A study explored the common features of high complexity complexity-aesthetics relationship for webpages using several past datasets. After accounting for the potential confounders – novelty and craftsmanship – we concluded that the novelty mainly affect the complexity while the craftsmanship mainly affect the aesthetics. If the craftsmanship was factored, the relationship between complexity and aesthetics appeared to be linear, a negative correlation. If the novelty was factored, there was no obvious relationship.

References

[1] Aliaksei Miniukovich and Maurizio Marchese. 2020. *Relationship Between Visual Complexity and Aesthetics of Webpages.* Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376602

[2] Litian Sun, T. Yamasaki, and K. Aizawa. 2014. Relationship Between Visual Complexity and Aesthetics: Application to Beauty Prediction of Photos. In *ECCV Workshops.*

[3] Alexandre N. Tuch, E. E. Presslaber, M. Stoecklin, K. Opwis, and Javier A. Bargas-Avila. 2012. The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. *Int. J. Hum. Comput. Stud.* 70 (2012), 794–811.

[4] Jian Wang and Yen Hsu. 2020. The Relationship of Symmetry, Complexity, and Shape in Mobile Interface Aesthetics, from an Emotional Perspective - A Case Study of the Smartwatch. *Symmetry* 12 (2020), 1403.