

Web Pages Aesthetic Evaluation Using Low-Level Visual Features

Maryam Mirdehghani, and S. Amirhassan Monadjemi

Abstract—Web sites are rapidly becoming the preferred media choice for our daily works such as information search, company presentation, shopping, and so on. At the same time, we live in a period where visual appearances play an increasingly important role in our daily life. In spite of designers' effort to develop a web site which be both user-friendly and attractive, it would be difficult to ensure the outcome's aesthetic quality, since the visual appearance is a matter of an individual self perception and opinion. In this study, it is attempted to develop an automatic system for web pages aesthetic evaluation which are the building blocks of web sites. Based on the image processing techniques and artificial neural networks, the proposed method would be able to categorize the input web page according to its visual appearance and aesthetic quality. The employed features are multiscale/multidirectional textural and perceptual color properties of the web pages, fed to perceptron ANN which has been trained as the evaluator. The method is tested using university web sites and the results suggested that it would perform well in the web page aesthetic evaluation tasks with around 90% correct categorization.

Keywords—Web Page Design, Web Page Aesthetic, Color Spaces, Texture, Neural Networks

I. INTRODUCTION

NOWADAYS, World Wide Web (WWW) plays a significant role in the information society and web page is the most important/unique interface between the user and that world. At the same time, we live in a period where visual appearances play an increasingly important role in our daily life. The web designer can never neglect the importance of visual appearance of their products in such a competitive market.

Since the proliferation of web sites began in the early 1990s, the relationship between usability and aesthetics has been a topic of heated discussions [1]. On one hand, usability experts believe on creating web sites where the user quickly and effectively can obtain the desired pieces of information without being delayed by long downloading times or complicated interfaces. On the other hand, there are graphic designers who believe that the web page should be visually interesting and aesthetically appealing.

In the second half of the 1990s, along with the widening bandwidth and increasing number of users, a renewed discussion that focused on the aesthetic effects in the web design has appeared. Now, design principles are going towards simple, usable, and aesthetic web pages [2]. Designers have realized that aesthetic arouses user's attention, engages the user to use the system, and helps to maintain his/her interest.

Up to now, there are several researches on the usability and principles of designing of usable web sites, e.g. [3]–[4]–[5]. Also, there are several attempts to design software which can analyses web page's components and place them properly to increase its usability [6]. Subject of aesthetic and its relationship to other design dimensions, is however very fresh.

In 1995, Kurosu and Kashimura conducted one of the first experiments ever to study correlations between user's aesthetic perceptions and their prior perceptions of the web page usability [7]. The results basically indicated that people expects that things that look good, actually perform better too. In 1997, Tractinsky duplicated the test and adapted it for different cultural settings and recorded an even stronger correlation between perception of aesthetic and apparent usability [8]. Since then, Tractinsky and some of his colleagues have conducted extensive research on the role that aesthetic plays in the information society [9]. In 2000, Chek, Ngo and Seng propose a theoretical approach to compute the essence of artists' insights with fourteen aesthetic measures for graphic design [10]. In 2004, Lavie and Tractinsky conducted studies to develop a measurement instrument for perceived web sites aesthetic evaluation, to be used in the future studies [11]. At the same time, Park and Choi attempted to identify critical factors that are closely related to the aesthetic fidelity of web pages [12]. Previously in 2002, Ivory and Hearts tried to create interactive tools to help non professional web site builders to create high quality designs. They computed a set of 157 measures over a large collection of pages [13]. In 2007, Beard published a book about principles of designing beautiful web design such as color, balance, texture, and layout and so on [14]. Previously in 2002, Lynch and Horton conducted studies on general principles for designing web pages and role of each designing component [15].

In 2007, Lindgaard *et al* show that the first impression assessed by visual appeal is very important [16]. They proof that "When surfers reach a web page, assess the visual appeal in about 50ms". This view helps them to decide if the page is relevant to their task and move towards the part of the page that interests them. Therefore, there should exist some especial features which a user implicitly employs to

S. Amirhassan Monadjemi is now an assistant professor at the Department of Computer Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran. (E-mail: monadjemi@eng.ui.ac.ir).

Maryam Mirdehghani is now an MSc student in AI and Robotics at the Department of Computer Engineering, Faculty of Engineering, University of Isfahan. (E-mail: mirdehghani@comp.eng.ui.ac.ir)

make decision about a web page quickly.

Meanwhile, visual features that affect user's decision making, varies on the basis of page's theme. Each theme has its own special viewers, so each page should attract most of its relevant audiences. This means that probably we can not find global features for all audiences (relevant or irrelevant) of a special theme. Therefore, we limit our case study and focus on the university web pages and their typical audiences which are students.

The main goal of this research is to train an intelligent system to perceive and quantify the first impression assessed by the web page's aesthetic, and based on this quantification, classify the web page into three groups: low quality (LQ), moderate (M) and high quality (HQ). We firstly develop an online questionnaire to collect people's perception and opinions about a known set of university web pages. Then we select some features which are supposed to be essential in designing beautiful web pages and tried to extract them from our web pages dataset using image processing techniques. When the feature vectors get ready, we feed them to an Artificial Neural Networks (ANN) with appropriate structure as the classifier, and see the ANN's opinion about the web page's aesthetic. At the following, we firstly describe the type of features we used, our classifier, and ends with the explanation and discussion of the experiments and results.

II. FEATURES

There are some basic visual features that are expected to be more significant than others in the aesthetic [14]–[17]. We consider the texture and color as two of those basic features. In the following section we describe each feature, its significance, and its extraction scheme separately.

A. Color Spaces

Visual design factors such as colors can strongly affect the viewers at the first glance. If we choose a vibrant, warm color for the menu, we communicate something different than if we had chosen a calm, cool blue scheme color. So we can use color features as a key point in aesthetic [14]. In this particular study, it is of importance to perceptual, human-like color features. Therefore, we tried HLS and CIE-Lab color spaces [17]. As will be mentioned later, HLS color features were more promising in our experiments. Employed color features were mostly hue and saturation of pixels or regions within the web pages, or some functions of them.

B. Gabor Filters

If we want to analyze the visual appearance of a web page, we should know the dependency amongst different components (here: pixels) of that. It will hint us to use the texture features of the web pages as another key aesthetic parameter. Texture features, in fact, show the correlation between pixels in a given neighborhood of the image [18].

There are different methods for analyzing textures, but one which perhaps fits the best to our study is multi-scale/multi-directional schemes, e.g. Gabor filtering. Using a proper Gabor Filter bank, the result of filtering will be a set of sub images, where each of them emphasizes on a certain frequency band and direction. So, we need to select appropriate set of directions and frequencies for our filters to

Vol:3, No:4, 2009 extract proper features from the image.

Equation (1) shows a Gabor filter in the frequency domain:

$$g(u, v) = \frac{1}{ab} \exp\left(-\pi\left(\frac{(u-u_0)^2}{a^2} + \frac{(v-v_0)^2}{b^2}\right)\right) \quad (1)$$

The parameters u_0 and v_0 define the spatial frequency of the filter center in a Cartesian coordinates and computed as $u_0 = F_0 \cos \omega_0$ and $v_0 = F_0 \sin \omega_0$, where ω_0 is the direction and F_0 is the central frequency of filter. The a and b are scaling parameters which are selected in a manner that neighbor filters in a filter bank catch each other in the half of their domain. An example of a Gabor filter bank is shown in Fig. 1.

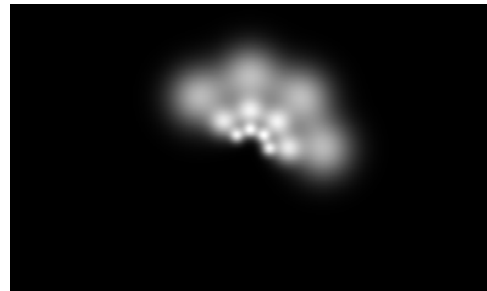


Fig. 1 Top view of a filters' bank consists of 12 filters with 4 direction $\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\}$ and 3 central frequency $\{\frac{f_m}{2}, \frac{f_m}{4}, \frac{f_m}{8}\}$, where f_m is the maximum frequency of image.

III. CLASSIFIER

Our final goal is to train a classifier which can classify new instances (i.e. web pages) on the basis of features extracted from known instances. This description is similar to the main applications of artificial neural networks.

Neural computing systems are adept at many pattern recognition task, more so than both traditional statistical and expert systems. Many different neural network architectures exist, each having their strength and limitations. In this study we considered a layered feed forward structure i.e. a Perceptron. The first or input layer serves as a holding site for the inputs applied to the network. The last or output layer is the point at which the overall mapping of the network input is available. Between these two extremes lies one layer of hidden neurons: it is in these internal layers that additional mapping or computing takes place [19].

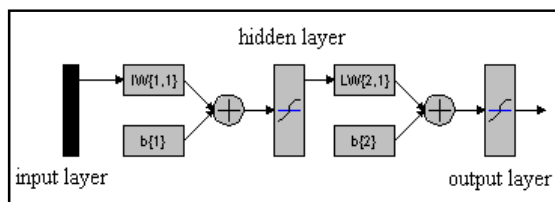


Fig. 2 Schematic view of an ANN

Links or weights connect each neuron in one layer only to those in the next higher layer. There is an implied directionality in these connections, in that the output of a neuron, scaled by the value of a connecting weight and squashing function, is fed forward to provide a portion of the activation for the neurons in the next higher layer.

Fig. 2 is a schematic representation of our ANN. We used a Feed forward network with 3 layers (one input, one hidden, and one output layer). Hidden layer uses tangent sigmoid as its squashing function, whereas the output layer uses a squashing linear function. This network is trained using back propagation method and its weights (LW) and bias (b) is determined. We used the neural network toolbox of MATLAB in our implementation.

IV. EXPERIMENTS AND RESULTS

As mentioned before, the main goal of this study is to develop an automatic web site aesthetic ranking system. To evaluate the methods, we firstly needed a reliable ground truth. We focused on the university web sites and in particular their first (main) pages. The first pages of 163 university web sites all around the world were captured and saved similarly. Next, as shown in Fig. 3, an online questionnaire was developed where the attendances were presented with 20 randomly selected web pages and asked to give each a 0 (very LQ) to 10 (very HQ) score. Around 210 students were attended this voting test in aggregation. We employed this ground truth for both testing and training the classifier. The average of scores would be computed for each 1000×600 web page image and put it into one of the three groups of {HQ, M, LQ} according to the TABLE I. The thresholds are in fact $\mu \pm \sigma$ of the average scores, assumed to have normal distribution.

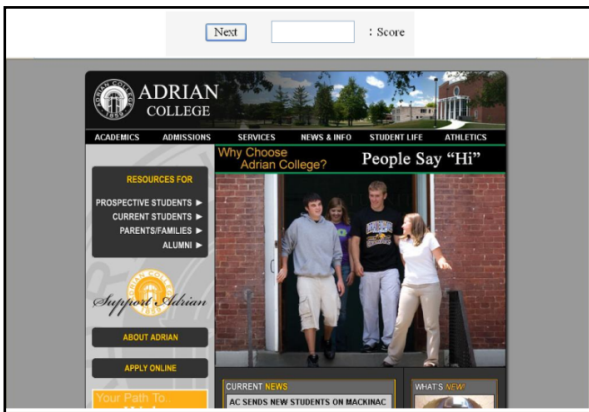


Fig. 3 Online Questionnaire Form and a test University web page.

TABLE I
AVERAGE SCORES AND CORRESPONDING GROUPS

Score (S)	Group
$S \geq 6.83$	HQ
$4.98 < S < 6.83$	M
$S \leq 4.98$	LQ

To extract texture features, we used Gabor filtering. As described earlier, each Gabor filter emphasizes on a specific direction and frequency band (i.e. scale). For constructing a proper Gabor filter bank, we considered two different sets of directions and three different sets of frequencies. Therefore, shown in TABLE II, we constructed 6 filter banks based on all possible combinations of those sets. The global and proportional energy of each sub image (i.e. the filter response) was considered as the feature. Using FB1, Fig. 4 illustrates the sub images of a given web page.

Color features employed in this study are based on the

used CIE-Lab color spaces. Firstly we reduced the resolution and number of colors of the image together in 8 consequent steps and 8 sub images were produced (see Fig. 5). Next, to extract HLS-based color feature, each sub image was converted to HLS and mean and standard deviation of H and S channels was computed separately. CIE-Lab feature vectors were built up by the same manner too. The results of our tests shows that HLS color space best reflex the peoples' perception of aesthetic. So, in all remaining tests we use HLS as our basic perceptual color space.



Fig. 4 Sub images (filter responses) of a web page using FB1 Gabor filter bank (see TABLE II). Rows are different central frequencies from low to high, and columns are directions.

TABLE II
GABOR FILTERS BANKS

Filter Bank#	Direction (degree)	Central Frequency (ω_c)
FB1	0, 45, 90, 135	1/2 1/4 1/8
FB2	0, 45, 90, 135	1/2 1/4 1/8 1/16
FB3	0, 45, 90, 135	1/2 1/4 1/8 1/16 1/32
FB4	0, 35, 70, 105, 140, 175	1/2 1/4 1/8
FB5	0, 35, 70, 105, 140, 175	1/2 1/4 1/8 1/16
FB6	0, 35, 70, 105, 140, 175	1/2 1/4 1/8 1/16 1/32

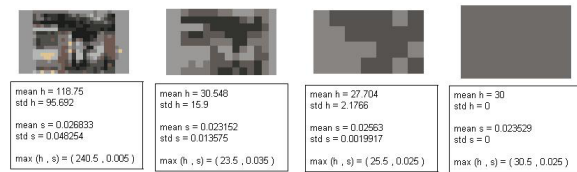
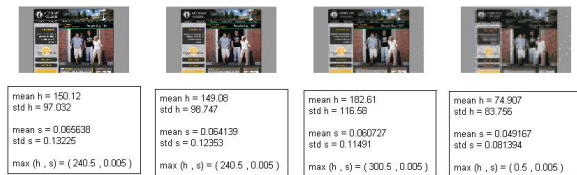


Fig. 5 Color features extraction, reducing the resolution along with the number of colors in 8 consequent steps.

At this point, the feature vectors are ready and we are going to use a multilayer feed forward Perceptron with 1 hidden layer as the classifier. The number of hidden neurons, which can strongly affect our results, along with the number of training epochs were selected during a trial and error procedure. In each test, around 80 web pages were selected as the training set including samples of all three categories. For testing, a non-overlapping set of 20 web pages was employed. The testing procedure was repeated 3 times for each trained network, and the average performance

was registered. Figure 4 illustrates the online questionnaire form used, and a test university first web page.

We firstly started our test using only the Gabor-based texture features. After series of optimization on both feature sets and the classifier, the correct classification performance was limited to 61.11% which was not promising. Next we tried the HLS-based color features which performed better with 72.22% correct classification at most. Our efforts to combine the texture and color features in a single feature vector were not successful, since the classification accuracy was also limited to 74.04%, mostly due to different nature of the features which decreases the homogeneity of the combined feature vector. Then a multi-classifier approach was tried, where each texture and color feature has got its own individual ANN classifier, but the overall score of the web page was computed by linear combination of both ANNs' output. As TABLE III depicts, the best performance achieved was 88.88% using $0.6y_{color}+0.4y_{gabor}$ linear combination. Again, the optimum combination coefficients were found by trial and error.

TABLE III
BEST LINEAR COMBINATION RESULTS OF TWO FEATURES

Filter Bank#	Filter Response Type	Combination Coefficients	Correct Decision (%)	Hidden Neurons#
FB1	Proportional frequency Std.	$0.6y_{color}+0.4y_{gabor}$	83.33%	10 – 20
FB1	Proportional Max. direction	$0.7y_{color}+0.3y_{gabor}$	88.88%	10 – 20

V. CONCLUSION

The main goal of this study is to develop a method to perceive impressions caused by aesthetic and classify web pages according to this perception. Although this perception is personal and differ from one person to another and depends on the web site theme, the study outcomes suggest that there are some special parameters which are common amongst the particular users (e.g. students) which are typical audiences of particular web pages (e.g. universities'). Also we have shown that ANN as a classifier can find the hidden aesthetic patterns behind the input features extracted from the texture and color of a web page.

Although the results of almost 90% correct aesthetic classification is promising, there are some limitations in this study. In particular, we limited our case study to around 200 students and 160 web pages. Again, more research is required to find out the most robust set of visual features for aesthetic evaluation. The possible advantages of the higher level image analysis should not be ignored too.

REFERENCES

- [1] Thorlacius, L.: The Role of Aesthetics in Web Design. In: Nordicom Review 28. pp. 63--76 (2007)
- [2] Lee, C.: HCI & Aesthetic: The Future of User Interface Design, <http://www.carrielee.net>.
- [3] Neilson, J.: Usability Engineering. Boston, San Diego, New York: Academic Press. (1993)
- [4] Neilson, J.: Designing Web Usability. Indianapolis. New Riders. (1999).
- [5] Ivory, M.Y., Sinha, R. R., Hearst, M. A.: Empirically Validated Web Page Design Metrics. In: ACM SIGCHI. pp. 53--60 (2001)

- [6] Gorman, B. Chimera, R. Jog, N. Stimart, R. White, D.: Evaluating spatial and textual style of displays. Technical report UMCP-CSD CAR-TR-763, CS-TR-3451, ISR-TR-95-51. (1995)
- [7] Kurosu, M., Kashimura, K.: Apparent Usability vs. inherent usability: experimental analysis on the determinants of the apparent usability. In: CHI 95. pp. 292--293 (1995)
- [8] Tractinsky, N.: Aesthetic and Apparent usability: Empirically Assessing Cultural and Methodological Issues. In: CHI 97. pp. 115--122 (1997)
- [9] Tractinsky, N.: Toward the Study of Aesthetic in Information Technology. In: 25th Annual International Conference in Information Systems, Washington DC, December 12-15, pp. 771--780 (2004)
- [10] Chek, D. Ngo, L. Seng Teo, L.: A Mathematical Theory of Interface Aesthetics. In: Visual Mathematics 2(No. 4), (2000)
- [11] Lavie, T., Tractinsky, N.: Assessing Dimension of Perceived Visual Aesthetic of Web Sites. In: International Journal of Human-Computer Studies, 60(3). pp. 269--298 (2004)
- [12] Park, S., Choi, D. Kim, J.: Critical Factors for the Aesthetic Fidelity of Web Pages: Empirical Studies with Professional Web Designers and Users. In: Interacting with Computers 16. pp. 351--376 (2004)
- [13] Ivory, M. Hearst, M.: Statistical Profiles of Highly-Rated Web Sites. In: ACM CHI 02, ISBN: 1-58113-453-3, pp. 367--374 (2002)
- [14] Beard, J.: The Principles of Beautiful Web Design. SitePoint, Pty. Ltd ISBN: 0-9758419-6-3 (2007)
- [15] Lynch, P. Horton, S.: Web Style Guide, <http://www.webstyleguide.com>.
- [16] Lindgaard, G. Fernandes, G. Dudekx, C. Brown, J.: Attention Web Designers: You Have 50 Milliseconds to Make a Good First Impression!. In: Behavior & Information Technology, Vol. 25, No. 2. pp. 115--126 (2006)
- [17] Gonzalez, R. C. Woods, R. E.: Digital Image Processing. Prentice Hall (2001)
- [18] Monadjemi, A.: Towards Efficient Texture Classification and Abnormality Detection. PhD thesis, University of Bristol, UK. (2004)
- [19] Schalkoff, Robert J.: Artificial Neural Networks, McGraw-Hill, ISBN:007057118 (1997)

Seyed Amirhassan Monadjemi, born 1968, in Isfahan, Iran. He got his PhD in computer engineering, pattern recognition and image processing, from University of Bristol, Bristol, England, in 2004. He is now working as a lecturer at the Department of Computer, University of Isfahan, Isfahan, Iran. His research interests include pattern recognition, image processing, human/machine analogy, and physical detection and elimination of viruses.

Maryam Mirdehghani, born 1982, in Tehran, Iran. She got his BSc in Computer Engineering, software, from Department of Computer Eng., Faculty of Engineering, University of Isfahan in 2005. She is now an MSc Student in Computer, AI and Robotics at the same academy. Her focus of study is artificial intelligence, image processing, and neural networks.