Learning to Evaluate the Visual Quality of Web Pages

Ou Wu
Institute of Automation
Chinese Academy of Sciences
wuou@nlpr.ia.ac.cn

Yunfei Chen, Bing Li
Institute of Automation
Chinese Academy of Sciences
{yfchen,bli}@nlpr.ia.ac.cn

Weiming Hu
Institute of Automation
Chinese Academy of Sciences
wmhu@nlpr.ia.ac.cn

ABSTRACT

A beautiful and well-laid out Web page greatly facilitates users' accessing and enhances browsing experiences. We use "visual quality (VQ)" to denote the aesthetics of Web pages. In this paper, a computational aesthetics approach is proposed to learn the evaluation model for the visual quality of Web pages. First, a Web page layout extraction algorithm (V-LBE) is introduced to partition a Web page into major layout blocks. Then, regarding a Web page as a semi-structured image, features known to significantly affect the visual quality of a Web page are extracted to construct a feature vector. The experimental results show the initial success of our approach. Potential applications include Web search and Web design.

Categories and Subject Descriptors

H.4.m [**Information Systems**]: Miscellaneous; H.2.8 [**Database Applications**]: Data Mining

General Terms

Algorithms

Keywords

Visual quality, aesthetics, semi-structured image, learning.

1. INTRODUCTION

Researchers in multiple disciplines have laid emphasis on the aesthetics of web pages. Findings on human computer interaction (HCI) suggest that aesthetics enhances positive feelings toward Web-based applications and has important implications for user experience [5]. Consequently, both the HCI and design research fields have exerted great efforts to improve a Web page's aesthetics and to develop standards to evaluate the aesthetics of a Web page [2]. Plenty of useful Web page design rules are established. However, due to limitations in visual information processing and Web mining techniques, previous studies have usually considered only a small number of factors. The lack of solid machine learning theories has resulted in that evaluation models are usually simple and have poor generalization capabilities. On the other hand, current Web mining mainly engages in the exploration of valuable content on the Web and in the evaluation of content usability. Web pages' aesthetics has receives litter attention in the Web mining literature despite its crucial effect on Web-based applications. To advance Web aesthetics research and construct a generalized, automatically and more accurate aesthetic evaluation model, this paper bridges the gap between the studies in the HCI and design communities and the methodologies in Web mining and machine learning. In this work, aesthetics is called visual quality (VQ). VQ evaluation can play a significant role in many Web-based applications. It can help Web page design and be intergraded in Web search.

WWW 2010, April 26–30, 2010, Raleigh, North Carolina, USA.
ACM 978-1-60558-799-8/10/04.

2. METHODOLOGY

Learning the VQ evaluation model for Web pages in our study involves a computational aesthetics approach. This falls into a classical statistical learning framework: extracting discriminative features and then learning the inductive model (e.g., classifier). However, VQ is subjective, which begs the question: is it possible to learn a model in order to measure the VQ of Web pages? Indeed, several pilot works [3] on aesthetic modeling for images have demonstrated encouraging results. The achievements on image aesthetic modeling suggest that the evaluation of the VQ of Web pages can be reasonably learned.

2.1 Overview of the proposed approach

A Web page can be viewed as a semi-structured image. The structural information includes the page layout, text positions and distributions, inner image positions, and background areas. We first apply a Web structure analysis tool to segment a Web page into a set of blocks, and then construct the layout for the page. We extract four classes of visual features for a page: layout visual features, text visual features, classical visual features (color and texture), and the visual complexity feature. To reduce the negative effects of a single person' s labeling subjectivity, each training page in our experiments is repeatedly labeled by several (seven in the experiments) persons. Finally, a cost-sensitive learning based classifier is constructed to classify Web pages into low-VQ or high-VQ category.

2.2 Page layout extraction

Michailidou et al. [4] concluded that there is a strong and high correlation between Web pages' layouts and their aesthetics. We introduce a heuristic layout extraction algorithm based on the Web page segmentation results. Given that this study explores the visual aspects of Web pages, the Vision-based Page Segmentation (VIPS) [1] algorithm is chosen as the basic segmentation algorithm. Our algorithm is called **V**IPS based **L**ayout **B**lock **E**xtraction algorithm (V-LBE). V-LBE first selects all the layout block candidates whose sizes are above a threshold (τ_1) and then deletes or inserts blocks to construct a set of un-overlapped large blocks which (approximately) cover the whole page. In our experiments, τ_1 is set as 1/9 of the whole page size, while τ_2 is set as 1/36 of the whole page size. Once the layout blocks are obtained, the number of major blocks on a page and the blocks' relative positions can be easily inferred. We use adjacent matrix (A) to describe the relationships between blocks: $A_{ij} = 1$ if the block i and the block j are adjacent, while $A_{ij} = 0$ otherwise.

2.3 Features

A Web page's VQ is influenced by a wide range of factors. With the extracted layout, a Web page can be transformed into a semi-structured image, and each block of the page corresponds to an image block. We consider four classes of features: (1) layout visual features (LV) that reflect the visual perception on the page layout; (2) text visual features (TV) that reflect the visual perception on the Web texts; (3) classical visual features including traditional color and texture features; and (4) the visual complexity feature (VC). The four classes of features characterize a Web page's four main parts: text, images, backgrounds, and the layout. Due to lack of space, we only give a sketch of each types of features. The detailed definitions can be referred to the full version of the paper. Thirty features are extracted as candidate features denoted as $\{f_i|1 \leq i \leq 30\}$. $f_1 - f_5$ characterize the layout visual features (**LV**), i.e.,the number of layout blocks (f_1), the sum of the width (f_2), aspect ratio (f_3), the number of leaf nodes (f_4), and the number of layers of the block tree (f_5). $f_6 - f_8$ describe the text visual features (**TV**) including text leaf node count (f_6), text area proportion (f_7) and character density (f_8). $f_9 - f_{20}$ characterize the color present in a page. Four color properties of a page are considered: Hue (**H**), Brightness (**Bri**), Saturation (**Sat**) and Colorfulness (**Col**). For each property, three features are extracted to reflect the average value, block contrast and block variance. $f_{21} - f_{29}$ represent the texture features (**T**) of a page. The ratio of the JPEG size to the whole Web-page image's area (f_{30}, **VC**) is alto considered because it reflects the visual complexity.

3. EVALUATION

We chose homepages as our experimental data and collected 500 homepages, mainly from sites of companies, universities, personal, and so on. To ensure that the gathered pages contain both high and low VQ pages, two persons collected high-VQ pages in their views, and two other persons collected low-VQ pages in their views. After the download, each page's layout and text blocks are extracted, and each page is transformed into a Web-page image. Each page was labeled by seven participants. Each participant is allowed to view one page within 5 seconds and assess the page from the five rating scores (-2, -1, 0, 1, and 2). "2" means very good, while "-2" means very bad. After human labeling, each page has seven scores. The average score is calculated. Its sign is as the label and its absolute value is the cost.

In cost-sensitive learning, mis-classified cost is applied to evaluate learning performances [6]. The average mis-classified cost (AMC) is used in this study.

Table 1: The classification performances (AMCs) of feature subset combinations

Combinations	AMC
LV, TV	0.2104
LV, TV, T	0.2084
LV, TV, T, VC	0.2063
LV, TV, T, VC, Col	0.1978
LV, TV, T, VC, Col, Bri	**0.1846**
LV, TV, T, VC, Col, Bri, Sat	0.2060
All features	0.2088

Cost-sensitive SVM [1] is used as the classification model with the radius-based function kernel. The AMCs achieved by leave-one-out validation are reported. The parameter C and the kernel parameter g are searched in {0.01, 0.1, 1, 10, 100} and {0.25, 0.5, 1, 2, 4, 8}, respectively.

We first rank the performances of each feature subset. Results show that layout visual features (LV) provide the best performance, and text visual features (TV) achieve the second-best performance.We undertake a forward feature selection strategy: the feature subsets are combined according to their performance orders. The AMCs of each feature subset combination are listed in Tbl. 1. It can be observed that (1) the best performance (AMC = 0.1846) is achieved by the feature combination of LV, TV, T, VC, Col, and Bri; (2) each feature subset combination outperforms all the single feature subsets; and (3) with the increase in the number of combined feature subsets, the AMC decreases at first and then increases. The best classification performance is close to the two labelers' AMC values (0.1422 and 0.1155).

4. CONCLUSIONS

In this paper, we have brought together multi-discipline studies from HCI, Web mining, machine learning, and computer vision to learn the VQ evaluation model for Web pages. To the best of our knowledge, this is the first work dedicated to the topic in the Web mining and machine learning literature. We have conducted experiments to evaluate the performances of different feature subsets and their combinations on the collected Web pages. The experimental results also indicate the promising of the proposed approach.

5. REFERENCES

[1] D. Cai. et al. vips: a vision-based page segmentation algorithm. *Microsoft Technical Report*, (MSR-TR-2003-79), 2003.

[2] P. Costa. Evaluating web site design. *In Proc. ACM SIGDOC'08*, pages 1071–1083, 2008.

[3] R. Datta. et al. studying aesthetics in photographic images using a computational approach. *In Proc. ECCV, LNCS*, pages 288–301, 2006.

[4] E. Michailidou. et al. visual complexity and aesthetic perception of web pages. *In Proc. SIGDOC'08*, pages 215–223, 2008.

[5] B. Schenkman and F. Jonsson. Aesthetics and preferences of web pages. *Behaviour and Information Technology*, 19(5):367–377, 2000.

[6] B. Zadrozny. et al. cost-sensitive learning by cost-proportionate example weighting. *In Proc. ICDM*, pages 435–442, 2003.

[1] http://svmlight.joachims.org/