# Aesthetic Visual Quality Evaluation of Chinese Handwritings

**Rongju Sun, Zhouhui Lian***, **Yingmin Tang, Jianguo Xiao**
Institute of Computer Science and Technology, Peking University, Beijing, P.R.China
{sunrongju, lianzhouhui, tangyingmin, xiaojianguo}@pku.edu.cn

## Abstract

Aesthetic evaluation of Chinese calligraphy is one of the most challenging tasks in Artificial Intelligence. This paper attempts to solve this problem by proposing a number of aesthetic feature representations and feeding them into Artificial Neural Networks. Specifically, 22 global shape features are presented to describe a given handwritten Chinese character from different aspects according to classical calligraphic rules, and a new 10-dimensional feature vector is introduced to represent the component layout information using sparse coding. Moreover, a Chinese Handwriting Aesthetic Evaluation Database (CHAED) is also built by collecting 1000 Chinese handwriting images with diverse aesthetic qualities and inviting 33 subjects to evaluate the aesthetic quality for each calligraphic image. Finally, back propagation neural networks are constructed with the concatenation of the proposed features as input and then trained on our CHAED database for the aesthetic evaluation of Chinese calligraphy. Experimental results demonstrate that the proposed AI system provides a comparable performance with human evaluation. Through our experiments, we also compare the importance of each individual feature and reveal the relationship between our aesthetic features and the aesthetic perceptions of human beings.

## 1 Introduction

With the booming development of image acquisition and visual computing techniques, currently there exists a large amount of research work on computer cognition in the literature. Computer-aided aesthetic evaluation is one of the hottest topics in artificial intelligence. However, previous work mostly focused on the evaluation of photographic images [Aydin *et al.*, 2014; Marchesotti *et al.*, 2014], paintings [Li and Chen, 2009] or videos [Yang *et al.*, 2011] [Yeh *et al.*, 2013]. Attempts to automatically make aesthetic judgments on Chinese handwritings are rare but quite practical in some related research areas, such as Chinese handwriting synthesis [Elarian *et al.*, 2014] [Li *et al.*, 2014] and Chinese handwriting beautification [Shi *et al.*, 2014]. Furthermore, the aesthetic evaluation module could also be quite helpful for font designers and handwriting learners.

*Zhouhui Lian is the corresponding author.

Figure 1: Chinese handwriting samples of eight different characters with good (first row) and bad aesthetic visual quality (second row).

As shown in Figure 1, Chinese handwriting images are monochromatic and the aesthetic visual quality of a Chinese handwriting is closely related to the character it represents. Thus, most classical image feature extraction methods that are based on colors, textures and shapes cannot be adopted directly. How to extract semantic features of aesthetic attributes according to the handwriting beauty principles is a challenging task. To address this problem, Lai *et al.* [Lai *et al.*, 1997] summarized four quantitative interpretations of traditional Chinese handwriting beauty rules and derived 10 different beauty evaluation metrics. These features are general but lack of the ability to represent sophistication. Recently, some researchers utilized online writing devices to obtain the dynamic information of handwritings. For example, a Wintab system was employed in [Wen, 2008] and Tobitani *et al.* [Tobitani *et al.*, 2008] used a Horizon View Camera to reconstruct the writing process. Meanwhile, some offline image-based methods were also proposed since writing on electronic devices is usually not convenient for calligraphy amateurs. Han *et al.* [Han *et al.*, 2008] developed a calligraphy learning system based on three quantized image-based aesthetic features. However, the size of the feature set is too small to handle some subtle and complex cases. Recently, Wang *et al.* [Wang *et al.*, 2014] proposed a hierarchical evaluation approach by matching the shape of the whole character and each stroke to the corresponding standard calligraphic template. Limited by the single template, this method ignores that the styles of beautiful handwritings can be varied to some extents. One of the most relevant work against this paper is [Xu *et al.*, 2007], where Xu *et al.* extracted strokes from a given handwriting image and fed a back propagation neural network with the features of each individual stroke shape, stroke spatial layout and consistency of stroke styles. However, these stroke-based features are closely related to the stroke extraction results while Chinese handwritings sometimes appear so cursive that it is even hard for human to separate every stroke accurately.

In general, there exist several limitations in current methods. First, existing aesthetic features are not sufficient enough to comprehensively depict the perspective metrics of Chinese handwritings based on handwriting rules. Second, Chinese handwritings sometimes are so cursive that it is hard to separate every stroke accurately and the single template based method cannot appreciate beautiful handwritings of other styles. In addition, there is still no any standard Chinese handwriting aesthetic evaluation database available to compare different algorithms. To solve these problems, we make the following three major contributions in this paper. 1) We design a number of novel global features as a map from objectively computable values to perceptually meaningful Chinese handwriting aesthetic rules. 2) We extract components, which are defined as parts that rarely intersect with others, instead of strokes and choose 10 recognized beautiful fonts as templates to encode the component layout features by the sparse coding method. 3) We propose a relatively large-scale Chinese Handwriting Aesthetic Evaluation Database (CHAED), which is publicly available on our website[1]. Experimental results show that the proposed method can make a similar decision on the aesthetic visual quality of a Chinese handwriting as human beings. Moreover, we analyze the importance of every individual feature, which may reveal the relationship between the proposed aesthetic features and human beings' aesthetic perceptions.

## 2 Method Description

In this paper, we solve the problem of the aesthetic visual quality evaluation for Chinese handwritings in the following three steps. First, a set of aesthetic features including global features and component layout features are designed from the point of art. Second, a relatively large-scale database is built by collecting handwriting images and human evaluation results. Finally, BP neural networks are trained on the database to get the ability of aesthetic evaluation. The data collection and the system training procedures will be discussed in Section 3 and 4, respectively. In this section, we concentrate on the main challenge, namely designing a set of aesthetic features, which are computable and closely related to handwriting beauty principles.

For the sake of seeking balance between feature generality and sophistication, we design two different types of features: global features and component layout features, which are both derived from authoritative calligraphy books and professional calligraphers' empirical recommendations.

[1]http://www.icst.pku.edu.cn/zlian/chin-beauty-eval/

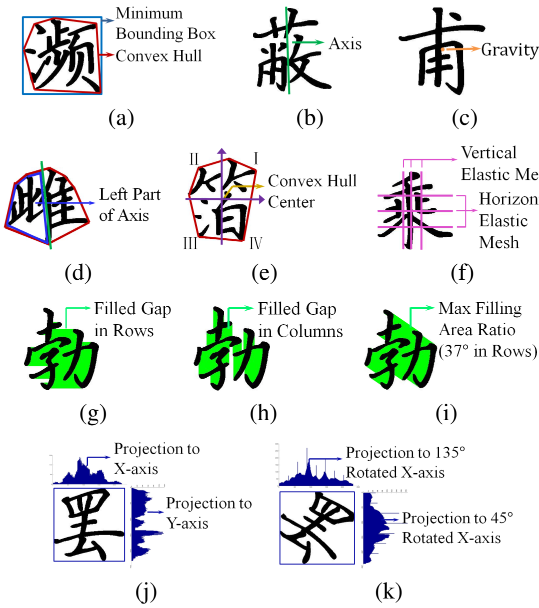



Figure 2: Demonstration of some global features.

Specifically, the proposed 22 global features are designed based on the overall impression of a handwritten Chinese character representing different aspects of the handwriting's aesthetic attributes. Moreover, components, which are defined as a set of strokes connecting interiorly but rarely overlapping with external strokes, are extracted using a semi-automatic method proposed in [Lian and Xiao, 2012]. Afterwards, we calculate a feature vector to describe the handwriting's layout structure in the component level and then utilize the sparse coding method [Wang *et al.*, 2010] to get 10 component layout features. Finally, each feature is normalized to the range $[0, 1]$ and all features are combined as a feature set $\Phi = \{f_i | 1 \leq i \leq 32, f_i \in [0, 1]\}$.

### 2.1 Global Features

According to our daily aesthetic experience, when people appreciate something, they first get a holistic impression and then go to details. Global features, which reflect the general visual impression towards a Chinese handwriting, are typically obtained from all pixels in the handwritten character image. Here, the proposed global features are formulated as a map from quantitative values to handwriting aesthetic rules. They are derived from the following three aesthetic aspects.

#### 1) Alignment and Stability

There exist various kinds of composition structures for different characters, such as the vertical composition structure, the horizontal composition structure, the surrounding composition structure, and so on. No matter which kind of structure the handwriting belongs to, it should have a good alignment to achieve a balanced structure in visualization. We design 3 kinds of features to describe the properties of alignment and stability as follows.

$f_1$**: Rectangularity of convex hull.** We employ the convex hull of a handwriting image, as shown in Figure 2(a), to represent the general shape of the handwritten character. If a stroke is far away from the central region of the character, the convex hull may have a sharp corner. Thus, the handwriting sample may be skew and oblique rather than aesthetically pleasing. The rectangularity of the character's convex hull can be calculated as $f_1 = \frac{P_c}{P_b}$, where $P_c$ and $P_b$ denote the perimeters of the convex hull and the minimum bounding box of the character, respectively. Obviously, $f_1$ ranges from 0 to 1. If the rectangularity value is equal to 1, the character's convex hull will become as a rectangle. Generally, the smaller the rectangularity value of convex hull is, the more irregular and unstable the handwriting is.

$f_2, f_3$**: Slope and intersection of axis.** Figure 2(b) shows the axis of a handwriting image, which divides all pixels into 2 subsets equally via a line: $y = kx + b$. The slope $k$ and the intersection $b$ are calculated by finding the coefficients of a polynomial prediction of degree 1 that fits the data best in a least-squares sense. The axis of a symmetrical and balanced handwritten character must be approximately perpendicular. We observe that the value of $k$ is in the range of 0 to 1, and thus we divide $b$ by $W_b$, which is defined as the width of the handwriting's minimum bounding box. Then, the two features can be computed by $f_2 = k$ and $f_3 = \frac{b}{W_b}$, respectively.

$f_4, f_5$**: Center of gravity.** In physics point of view, the mass center can be used to describe the stability of a rigid body. Given a handwriting image (e.g., Figure 2(c)), its center of gravity is calculated as $(x_g, y_g) = \frac{1}{C}\sum_{i=1}^{W_b}\sum_{j=1}^{H_b}(x_i, y_i) \times I_{i,j}$, where $C = \sum_{i=1}^{W_b}\sum_{j=1}^{H_b} I_{i,j}$ denotes the number of all black pixels in the image, while $W_b$ and $H_b$ are the width and height of the minimum bounding box, respectively. If $(i, j)$ is a black pixel, $I_{i,j} = 1$, otherwise $I_{i,j} = 0$. Since the center of gravity is located inside the character's minimum bounding box, we normalize these two features as $f_4 = \frac{x_g}{W_b}$ and $f_5 = \frac{y_g}{H_b}$.

**2) Distribution of White Space**

This kind of features represent the compactness of a Chinese handwriting indicating whether the black pixels in the image are crowded or not. The ratio of white space area is defined as White Space Ratio (WSR). If the WSR value of the handwriting image is small, that often means the handwriting is crowed. While if the value is close to 1, the handwriting may be too loose.

$f_6$**: Convexity.** A classic and simple method to measure the white space distribution is to get the proportion of black pixels in the image. Let $A_{convex}$ be the area of the convex hull, which is approximated as the total number of both the black and white pixels, and $C$ be the number of the black pixels, the convexity can be calculated as $f_6 = \frac{C}{A_{convex}}$.

$f_7$**: Ratio of axis cutting convex hull.** As mentioned above, the character's axis represents the distribution of pixels and the convex hull represents the general shape of the handwriting. Therefore, we can formulate the description of WSR as the proportion of area on the left part of the convex hull divided by the axis (see Figure 2(d)). Thus, we have $f_7 = \frac{A_{left}}{A_{convex}}$.

$f_8, f_9, f_{10}, f_{11}$**: Ratios of pixel distribution in quadrants.** As shown in Figure 2(e), let $(x_c, y_c)$ be the center of the character's convex hull and thus a new coordinate system can be established with $(x_c, y_c)$ as the origin. Then, the character image can be divided into the four quadrants. We calculate the ratio of the pixel quantity $C_i$ in the corresponding quadrant to the area of convex hull $A_{convex(i)}$ as follows: $f_{7+i} = \frac{C_i}{A_{convex(i)}}$, where $i = 1, 2, 3, 4$ denote Quadrant I, II, III and IV of the coordinate system, $C_i$ is the number of black pixels and $A_{convex(i)}$ is the area of the convex hull in the $i$th quadrant.

$f_{12}, f_{13}, f_{14}, f_{15}, f_{16}, f_{17}$**: Elastic mesh layout.** Elastic meshes are built by equally dividing black pixels of a given character image both in the horizontal and vertical directions. It has been considered as an important tool to recognize Chinese handwritten characters [Jin and Wei, 1998]. As shown in Figure 2(f), here we choose $4\times4$ meshes to partition the black pixels in the image equally. In order to obtain features whose values range from 0 to 1, the positions of vertical and horizontal separation lines should be divided by $W_b$ or $H_b$, respectively. Thus, we have $f_{11+i} = \frac{x_{vLine(i)}}{W_b}$, where $i = 1, 2, 3$ and $x_{vLine(i)}$ denotes the position of the $i$th vertical line, and $f_{14+j} = \frac{y_{hLine(j)}}{H_b}$, where $j = 1, 2, 3$ and $y_{mesh(j)}$ represents the position of the $j$th horizontal line.

**3) Gap Between Strokes**

Besides the white space distribution, we also expect to have features to represent the stroke's orientation and position. We extract the following two types of features, namely the maximum gap proportion and the variance of pixels' projection.

$f_{18}$**: Maximum gap proportion.** Inspired by the shape density analyzing method reported in [Qu *et al.*, 2014], we choose to fill the gap of a handwritten character in the vertical direction by drawing lines between the topmost black pixel and the bottommost black pixel in rows and columns of the image, as illustrated in Figure 2(g) and 2(h). These features are adopted to describe the distance between strokes. In order to eliminate the effect of difference between characters and achieve pose invariance, the gap is filled when rotating the handwriting image per degree and then the maximum filling area ratio can be obtained as shown in Figure 2(i). Thus, we have $f_{18} = \max\{\frac{C_{gap(\alpha)}}{C_{gap(\alpha)}+C}\}$, where $\alpha = 1°, 2°, ..., 90°$, which means rotating the image by the angle $\alpha$, and $C_{gap(\alpha)}$ is the number of pixels (see the green parts in Figure 2(g), 2(h) and 2(i)) filling the gap when the image rotates by $\alpha$.

$f_{19}, f_{20}, f_{21}, f_{22}$**: Variance of pixels' projection.** The four most frequently used types of strokes in Chinese characters are horizontal strokes, vertical strokes, left-falling strokes and right-falling strokes. The existence of these strokes can be observed by projecting black pixels to the $0°, 45°, 90°$ and $135°$ rotated x-axis, respectively. Figure 2(j) and 2(k) show that the character "ba" includes horizontal strokes because the handwriting image generates projection peaks in a small range in y-axis. The projection distribution can be described by the variance $\delta_\alpha$ and we denote the features as $f_{18+i} = \delta_\alpha$, where $i = 1, 2, 3, 4$ and $\alpha = 0°, 45°, 90°, 135°$, which represents the projection of black pixels to x-axis by rotating $\alpha$ degree.

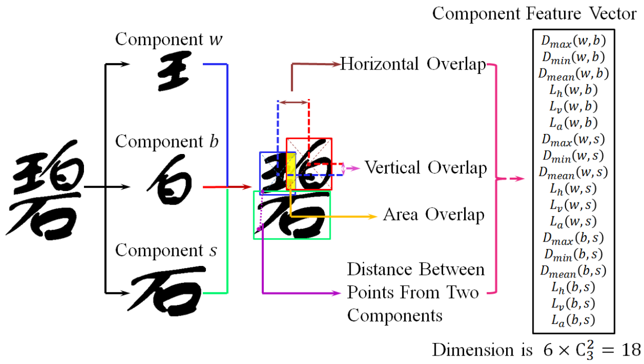



Figure 3: The handwritten character “bi” is decomposed into 3 components and the component feature vector is extracted to describe the components’ spacial layout.

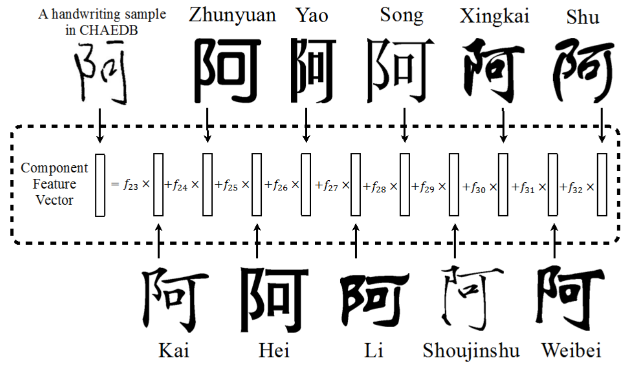



Figure 4: The handwriting’s component feature vector is coded by the component feature vectors of 10 template fonts.

### 2.2 Component Layout Features

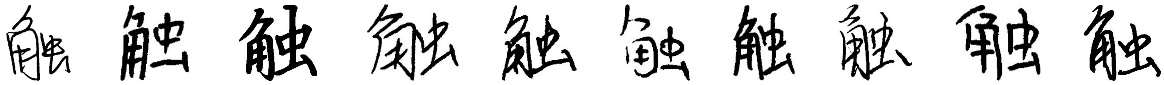
As important as global features, layout features are needed to depict the way how strokes or/and radicals are arranged to compose a character. Xu *et al.* [Xu *et al.*, 2007] extracted strokes to get the layout features. However, many handwriting characters are very cursive, and thus it is difficult to extract strokes and the extraction results could seriously affect the layout features. To solve this problem, we choose to use the specified components instead of strokes in our method. Here, the component is defined as a set of strokes, which are connected interiorly but rarely overlap with external strokes, such as $w$, $b$ and $s$ shown in Figure 3. Therefore, it is much easier to extract components from a handwriting image compared to strokes. Here, we adopt a semi-automatic component extraction method proposed in [Lian and Xiao, 2012]. Using this method, components can be automatically extracted for most characters and the decomposition results can be interactively corrected through a user interface for some very cursive handwritings when necessary.

After component extraction, we compute the maximum distance, minimum distance, and mean distance between every pair of components and divide these distances by the diagonal distance of the character’s minimum bounding box $\sqrt{W_b(r)^2 + H_b(r)^2}$ for normalization. Then we get $D_{max}(a,b)$, $D_{min}(a,b)$ and $D_{mean}(a,b)$ for the two components $(a,b)$. To reduce the computational complexity, contours are obtained from component images to compute these 3 features instead of using the original images. Furthermore, three types of overlapping ratios are computed including the horizontal overlapping ratio $L_h(a,b) = \frac{x_c(a)-x_c(b)}{W_b(a)+W_b(b)}$, the vertical overlapping ratio $L_v(a,b) = \frac{y_c(a)-y_c(b)}{H_b(a)+H_b(b)}$ and the area overlapping ratio $L_a(a,b) = \frac{A_b(a)\cap A_b(b)}{A_b(a)\cup A_b(b)}$, where $(x_c(r), y_c(r))$ denotes the center of the minimum bounding box of component $r$, $W_b(r)$ and $H_b(r)$ are the minimum bounding box’s width and height, and $A_b(r)$ denotes the area of the component’s minimum bounding box. All the above-mentioned features constitute a component layout feature vector and the dimension of the vector is $6 \times C_n^2$, where $n$ stands for the number of components in the character. Figure 3 demonstrates the construction process of the component feature vector.

Since the dimension of the component layout feature vector varies for different characters, it is hard to utilize them in a machine learning system. To address this problem, 10 recognized beautiful fonts are selected, namely Song, Kai, Hei, Weibei, Xingkai, Li, Fangsong, Shoujinshu, Yao and Zhunyuan styles. We first obtain the handwritten character image’s component feature vector $V$ and the corresponding 10 templates’ component feature vectors. Then, according to the Locality-constraint Linear Coding (LLC) [Wang *et al.*, 2010] method, we calculate the coefficients $c$ of the 10 templates by solving $\min \sum_{i=1}^{k} \|V - c_i B_i\|^2$, where the values of coefficients are constrained by $\sum_{i=1}^{k} c_i = 1$ and the codebook $B_i, i = 1, ..., k$ are the K-Nearest Neighbors of $V$. In this paper, we experimentally set $k = 5$. As shown in Figure 4, we get $f_{22+i}, i = 1, 2, ..., 10$ which denote the handwriting’s component layout features. As we know, some characters only consist of one component and thus their component layout feature vectors are empty. Therefore, when we refer to the component layout features, we only take the characters with multi-components into consideration.

## 3 Chinese Handwriting Aesthetic Evaluation Database

As we know, aesthetic evaluation is a subjective task, and thus Chinese handwriting aesthetic visual quality evaluation can be treated as a data-driven learning problem. However, since it is not easy to design and collect evaluation data, currently there is still no database publicly available for such purpose. To build a reliable and convincing database, we design CHAED based on a notable degree of agreement between aesthetic judgments of different people. Specifically, we select 100 characters, including 20 characters with single element and 80 with multi components in different structure types including horizontal composition, vertical composition, half surrounding composition and surrounding composition as shown in Figure 6. The size of each subset is determined by the diversity and complexity of its structure. 30 students are selected randomly to write the 100 characters. The images



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j)

Figure 5: 10 handwriting samples of Chinese character "chu" in CHAED. The left five images are from the training set while the right five images are from the testing set. Table 1 shows the scores each image received. The grading errors of our algorithm with hybrid features $E_o$ are smaller than human grading variance $V_h$ in general. Also, $E_o$ are generally smaller than the errors obtained using the other two feature sets, namely $E_g$ and $E_c$.

Table 1: Comparison of algorithmic and human evaluation scores for the handwriting samples in Figure 5.

| | | **a** | **b** | **c** | **d** | **e** | **f** | **g** | **h** | **i** | **j** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Human | $S_h$ | 43.75 | 21.87 | 95.45 | 32.81 | 45.31 | 18.75 | 31.25 | 75.00 | 17.18 | 76.56 |
| Evaluation | $V_h$ | **33.92** | **28.28** | **14.59** | **31.83** | **32.92** | **28.03** | **27.00** | **30.92** | **27.78** | **33.28** |
| Global Features | $S_g$ | 45.93 | 34.14 | 100.66 | 42.92 | 46.73 | 48.56 | 56.96 | 42.88 | 43.85 | 94.80 |
| | $E_g$ | 2.18 | 12.27 | 5.21 | 10.11 | 1.42 | 29.81 | 25.71 | 32.11 | 26.66 | 18.24 |
| Component | $S_c$ | 47.08 | 46.65 | 44.94 | 47.88 | 45.49 | 49.39 | 45.31 | 46.97 | 47.36 | 47.32 |
| Layout Features | $E_c$ | 3.33 | 24.77 | 50.51 | 15.07 | 0.17 | 30.64 | 14.06 | 28.02 | 30.17 | 29.24 |
| Hybrid Features | $S_o$ | 37.48 | 31.04 | 98.05 | 37.60 | 40.92 | 45.80 | 36.72 | 41.54 | 57.03 | 66.58 |
| | $E_o$ | **6.26** | **9.17** | **2.59** | **4.79** | **4.38** | **27.05** | **5.47** | **33.46** | **39.85** | **9.98** |

(a) (b) (c) (d) (e)

Figure 6: 5 Chinese handwriting samples with (a) single-element, (b) horizontal composition, (c) vertical composition, (d) half surrounding composition and (e) surrounding composition.

are scanned and preprocessed in the same way to eliminate the effect of variances in image quality. Among these images and a number of handwritten characters acquired from some professional calligraphers, we pick up 10 images for each character by the criteria of allocating good samples, medium samples and bad samples in balance.

To collect the aesthetic evaluation data, a website is built and 33 subjects are invited to grade the 1000 handwritings into 3 levels: good, medium and bad. Generally, the probabilities of good, medium and bad evaluation results of each image are denoted as $p_{good}$, $p_{medium}$ and $p_{bad}$. We define the aesthetic score by

$$S = 100 \times p_{good} + 50 \times p_{medium} + 0 \times p_{bad}, \quad (1)$$

where $p_{good}$, $p_{medium}$ and $p_{bad}$ are in the range of [0, 1]. Based on the Eq. 1, we can calculate the average human evaluation score $S_h$ for each image. Afterwards, the absolute grading error between each human evaluation score and its corresponding image's average human evaluation score is computed as $E_{h_i} = |S_i - S_h|$. Let $E_h$ be the set of human evaluation errors, the quadratic mean of $E_h$ belonging to a character constitutes $V_h = \sqrt{\frac{\sum_{i=1}^{33} {E_{h_i}}^2}{n}}$ which means the score variance of $S_h$. For global features, half of the database is used for training and the other half for testing. To be specific, for each character, there are 5 handwriting samples for training and 5 for testing. Note that, when evaluating the component layout features, we only extract those features for the 80 characters with multi-components.

## 4 Experiments

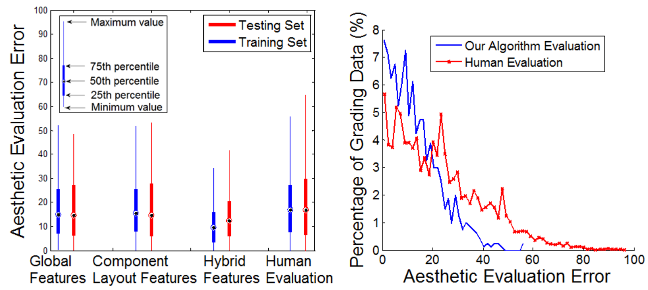
In this section, we carry out experiments on the CHAED database by using back-propagation (BP) neural networks with inputs of global features, component layout features and hybrid features (i.e., the combination of global and layout features), respectively, to evaluate the aesthetic quality of Chinese handwritings. Then, we compare the evaluation errors obtained by the proposed algorithms with human evaluation errors (see Figure 7(a) and 7(b)). Furthermore, the performance of each individual feature is also evaluated and we try to analyze several most important features to build a bridge between computer cognition and human perception.

### 4.1 Aesthetic Evaluation Performance

We build three 4-layer back-propagation neural networks denoted as $Net_g$, $Net_c$ and $Net_o$ for global features, component layout features and hybrid features, respectively. The networks are fed by the feature sequences of the training set with 500, 400 and 400 images, respectively. The sizes of training sets are different because we only take multi-component characters into account when using the component layout features. Thus, the input dimensions for the 3 networks are 22, 10 and 32, respectively, and the output is a probabilistic evaluation result in the form of $(p_{good}, p_{medium}, p_{bad})$, where $p_{good}$, $p_{medium}$ and $p_{bad}$ are probabilistic values of the 3 aesthetic levels. Then, the scalar score can be calculated based on the three probabilistic values by using Eq. 1. We determine the structure of the 3 neural networks by adjusting the training function, adaption learning function and the number of neurons in every layer to achieve the best evaluation results in the training dataset. Here, we choose TRANGDM as the training

Table 2: Top 4 most important global features for the aesthetic evaluation of Chinese handwritings.

| Feature | Meaning of Feature | Aesthetic Aspect | Accuracy |
|---|---|---|---|
| $f_6$ | Convexity of convex hull | Distribution of white space | 69% |
| $f_{19}$ | Variance of projection to x-axis | Gap between components | 60.8% |
| $f_2$ | Slope of axis | Alignment and stability | 60.6% |
| $f_{21}$ | Variance of projection to y-axis | Gap between components | 60.6% |




(a) (b)

Figure 7: Statistical distributions of the grading errors of our algorithm's grading results and human grading results in CHAED.

function and LEARNGDM as the adaption learning function for these 3 networks. The number of neurons for $Net_g$, $Net_c$ and $Net_o$ are respectively $(22, 20, 10, 3)$, $(10, 15, 10, 3)$ and $(32, 40, 20, 3)$. We assess the performance by the absolute difference between our algorithm's grading score $S_g$, $S_c$ and $S_o$ and the average human grading score $S_h$. The grading errors of the networks are calculated by $E_a = |S_a - S_h|$, where $a = g, c, o$ denotes the system using global features, component layout features and hybrid features, respectively. Grading errors of human evaluation results $E_h$ are taken as the ground truth, which indicates human's variance in appreciating handwritings, as shown in Figure 7(a). Moreover, distributions of evaluation errors using the hybrid feature and human grading errors are compared in Figure 7(b).

In Figure 7(a), the average value of $E_o$ is lower than that of $E_g$ and $E_c$, which indicates that the combination of global features and component layout features improves the evaluation performance significantly. Therefore, these two kinds of features are complementary. Furthermore, the average values of $E_g$, $E_c$ and $E_o$ are all not larger than that of $E_h$, which demonstrates that our algorithm can evaluate the aesthetic visual quality of Chinese handwritings as well as human beings. In order to compare the distribution of $E_o$ and $E_h$, we take the percentage of grading data instead of the number of grading data into account to draw the distribution curves in Figure 7(b). We can see that a larger proportion of data in $E_o$ is distributed in the range of [0, 20] and a smaller proportion in [20, 100] than $E_h$. Since $E_h$ represents human grading errors, it is reasonable to conclude that our system evaluates Chinese handwritings even better than many human beings. 10 handwriting samples of character "chu" and their evaluation results are shown in Figure 5 and Table 1.

### 4.2 Feature Importance Analysis

In addition to the overall aesthetic evaluation performance, we also want to know more about the relations between our features and human perceptions in appreciating Chinese handwritings. As we can see, the proposed global features have real aesthetic meanings, while the component layout features do not represent aesthetic attributes directly. Therefore, we choose to apply Support Vector Machine (SVM) to evaluate the performance for each individual global feature separately. More specifically, each global feature $f_i$ is chosen as the input and SVM is used to classify the handwriting images with the output of probabilities of 3 aesthetic levels $(p_{good}, p_{medium}, p_{bad})$. We choose the maximum value in $(p_{good}, p_{medium}, p_{bad})$ and take the corresponding aesthetic level as the handwriting's label. By comparing the computer-generated results with human evaluation results, labeled by choosing the maximum probability value as well, we can get the classification accuracies of each feature for all images in our database.

Experimental results show that the classification accuracy of each individual global feature is low in general but 4 features are obviously better than others. The top 4 features are listed in Table 2. We try to give a meaningful interpretation of their importance based on art knowledge and psychology theory.

$f_6$: Intuitively, the distribution of white and black areas affects people's impression on a Chinese handwriting directly. $f_6$ represents the proportion of black pixels in the character's core region, namely the convex hull. A balanced and harmonious handwriting must be neither too crowded nor too loose. Therefore, a given Chinese handwriting is very likely to be low-quality if the convexity of its convex hull is very large or very small.

$f_{19}, f_{21}$: High quality handwritings are more likely to be regular with straight horizontal and vertical strokes, which can be observed through the projection of black pixels to x-axis and y-axis. This is a unique characteristic for Chinese characters, which is quite distinctive against western languages. We may feel uncomfortable to some extent when facing a distorted and disordered handwriting. Hence, distributions of pixels projected in horizontal and vertical directions can be used to distinguish whether the Chinese handwriting can lead to a harmony and peace.

$f_2$: Chinese people typically appreciate symmetrical creatures, not only in architectures but also in handwritings. The slope of a handwriting's axis is comparable to the beam of a building. If the beam is oblique, the building seems unsafe. Similarly, a handwriting image with a lean axis may bring unstable and insecure feelings. This kind of feelings may cause slight anxieties instead of pleasing emotions.

## 5 Conclusion and Future Work

In this paper, we presented a framework to evaluate the aesthetic visual quality of Chinese handwritings. Specifically, we proposed 22 global features and 10 component layout features according to the art knowledge from books and calligraphers. Meanwhile, 1000 images of 100 characters were collected and 33 subjects were invited to grade the handwriting samples through a website. Then, BP neutral networks were trained to output the probabilities of the 3 aesthetic levels. Experiments show that our systems are able to evaluate the aesthetic quality of Chinese handwritings comparable to human beings. Furthermore, SVM was also utilized to analyze the importance of every individual global feature, which helps us to understand further about the aesthetic quality assessment by connecting with some art knowledge and psychology theory. In the future, we will keep our efforts on discovering semantic aesthetic features through cooperating with the Chinese calligraphy community. Also, CHAED will be enlarged by adding more handwriting images and collecting more evaluation data.

## Acknowledgments

This work was supported by National Natural Science Foundation of China (Grant No.: 61202230 and 61472015), National Hi-Tech Research and Development Program (863 Program) of China (Grant No.: 2014AA015102) and Beijing Natural Science Foundation (Grant No.: 4152022).

## References

[Aydin *et al.*, 2014] T Aydin, Aljoscha Smolic, and Markus Gross. Automated aesthetic analysis of photographic images. *Transactions on Visualization and Computer Graphics*, 21(1):31–42, 2014.

[Elarian *et al.*, 2014] Yousef Elarian, Radwan Abdel-Aal, Irfan Ahmad, Mohammad Tanvir Parvez, and Abdelmalek Zidouri. Handwriting synthesis: classifications and techniques. *International Journal on Document Analysis and Recognition*, 17(4):455–469, 2014.

[Han *et al.*, 2008] Chin-Chuan Han, Chih-Hsun Chou, and Chung-Shiou Wu. An interactive grading and learning system for chinese calligraphy. *Machine Vision and Applications*, 19(1):43–55, 2008.

[Jin and Wei, 1998] Lianwen Jin and Gang Wei. Handwritten chinese character recognition with directional decomposition cellular features. *Journal of Circuits, Systems, and Computers*, 8(04):517–524, 1998.

[Lai *et al.*, 1997] Pak-Keung Lai, Dit-Yan Yeung, and Man-Chi Pong. A heuristic search approach to chinese glyph generation using hierarchical character composition. *Computer Processing of Oriental Languages*, 10(3):281–297, 1997.

[Li and Chen, 2009] Congcong Li and Tsuhan Chen. Aesthetic visual quality assessment of paintings. *IEEE Journal of Selected Topics in Signal Processing*, 3(2):236–252, 2009.

[Li *et al.*, 2014] Wei Li, Yuping Song, and Changle Zhou. Computationally evaluating and synthesizing chinese calligraphy. *Neurocomputing*, 135:299–305, 2014.

[Lian and Xiao, 2012] Zhouhui Lian and Jianguo Xiao. Automatic shape morphing for chinese characters. In *SIGGRAPH Asia 2012 Technical Briefs*, page 2, 2012.

[Marchesotti *et al.*, 2014] Luca Marchesotti, Naila Murray, and Florent Perronnin. Discovering beautiful attributes for aesthetic image analysis. *International Journal of Computer Vision*, pages 1–21, 2014.

[Qu *et al.*, 2014] Jingwei Qu, Xiaoqing Lu, Lu Liu, Zhi Tang, and Yongtao Wang. A method of density analysis for chinese characters. In *Natural Language Processing and Chinese Computing*, pages 54–66. Springer, 2014.

[Shi *et al.*, 2014] Cao Shi, Jianguo Xiao, Canhui Xu, and Wenhua Jia. Nonlinear and non-gaussian bayesian based handwriting beautification. In *IS&T/SPIE Electronic Imaging*, pages 902012–902012, 2014.

[Tobitani *et al.*, 2008] Kensuke Tobitani, Nobuyasu Okabe, Kazuhiko Yamamoto, and Kunihito Kato. A proposal of evaluation method for balance of white space in calligraphy by using horizon view camera. In *Document Analysis Systems*, pages 463–469. IEEE, 2008.

[Wang *et al.*, 2010] Jinjun Wang, Jianchao Yang, Kai Yu, Fengjun Lv, Thomas Huang, and Yihong Gong. Locality-constrained linear coding for image classification. In *Computer Vision and Pattern Recognition*, pages 3360–3367, 2010.

[Wang *et al.*, 2014] Mengdi Wang, Qian Fu, Zhongke Wu, Xingce Wang, and Xia Zheng. A hierarchical evaluation approach of learning chinese calligraphy. *Journal of Computational Information Systems*, 10(18):8093–8107, 2014.

[Wen, 2008] Limin Wen. Construcion of an aesthetic features based calligraphy evaluation model. *Computer Engineering and Design*, 29(7):1865–1868, 2008.

[Xu *et al.*, 2007] Songhua Xu, Hao Jiang, Francis Chi-Moon Lau, and Yunhe Pan. An intelligent system for chinese calligraphy. In *Proceedings Of The National Conference On Artificial Intelligence*, volume 22, page 1578, 2007.

[Yang *et al.*, 2011] Chun-Yu Yang, Hsin-Ho Yeh, and Chu-Song Chen. Video aesthetic quality assessment by combining semantically independent and dependent features. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1165–1168, 2011.

[Yeh *et al.*, 2013] H Yeh, C Yang, M Lee, and C Chen. Video aesthetic quality assessment by temporal integration of photo-and motion-based features. *Transactions on Multimedia*, 15(8):1944–1957, 2013.