Learning Aesthetics Measure of a Document Page Layout From Designers' Interaction

Ildus Ahmadullin; Niranjan Damera-Venkata

HP Laboratories
HPL-2012-13

Keyword(s): automated publishing; quantifying aesthetics; designer interaction; probabilistic modeling

Abstract:
Aesthetics evaluation of a document layout is typically performed by a designer. Recently there has been proposed a number of automated systems for document creation. We present a new paradigm to automatically quantify document aesthetics, that can be used in the automated document composition techniques. Given a document template, we specify a few aesthetics parameters that can be adjusted to obtain the most aesthetically pleasing layout of a document page. We use graphical interface, where for a given amount of content designers can adjust the parameters to obtain a layout of the highest aesthetic value. The parameters are stored as a feature vector. After obtaining sufficiently large amount of feature vectors, the feature set is modelled by a Gaussian probability distribution. The resulting model can be used in predicting aesthetic value of a new document, or to compose a document of the highest aesthetic value for a given content.
Abstract

Aesthetics evaluation of a document layout is typically performed by a designer. Recently there has been a proposal of a number of automated systems for document creation. We present a new paradigm to automatically quantify document aesthetics, that can be used in the automated document composition techniques. Given a document template, we specify a few aesthetics parameters that can be adjusted to obtain the most aesthetically pleasing layout of a document page. We use graphical interface, where for a given amount of content designers can adjust the parameters to obtain a layout of the highest aesthetic value. The parameters are stored as a feature vector. After obtaining sufficiently large amount of feature vectors, the feature set is modelled by a Gaussian probability distribution. The resulting model can be used in predicting aesthetic value of a new document, or to compose a document of the highest aesthetic value for a given content.


Keywords: automated publishing, quantifying aesthetics, designer interaction, probabilistic modeling

1 Introduction

Aesthetic document composition is typically considered a hard problem, often requiring professional designer skills. In the traditional publishing industry producing hand-designed single page documents is widely practiced. However, due to high marginal cost of designer services, hand-designed creation of personalized documents is not economically viable. Therefore, automated document layout composition has been the topic that attracted a lot of research. [HURST et al. 2009]

One of the most important concerns in automated document composition systems is producing aesthetically appealing document layouts. It is unrealistic to assume that automatically designed documents can overcome performance of a professional designer. However, some reasonable means to measure aesthetics and distinguish a good layout from an obviously bad one is highly desirable.

One of the most popular works in this area is probably [HARRINGTON et al. 2004]. In this paper the authors study the attributes that degrade the aesthetic quality of a document. Non-linear combination of heuristic measures is proposed to predict the aesthetics of a document. It is claimed that the implementations are relatively efficient and further research in this approach is encouraged.

As opposed to deterministic methods of quantifying document aesthetics, the models can also be designed probabilistically. Each of the document element parameter can be assigned a probability distribution that signifies aesthetics. After combining such distributions one can come up with an overall probabilistic framework of determining aesthetic quality of the layout. [Damera-Venkata et al. 2011] assignes Gaussian prior probability distributions to aesthetics parameters. The distribution parameters are set heuristically. The algorithm uses a probabilistic graphical model to calculate optimal layout parameters, and the aesthetics prior distributions are playing major role in document layout composition.

All of these methods, however lack human designer input. The heuristics in measuring document aesthetics are based on observation of designer prepared layouts. In this paper, we utilize designers’ input directly in our algorithm. Once we train the algorithm for a certain template, it can mimic the designers’ evaluation in measuring layout aesthetics.

2 Learning from designers’ interaction

Document layout can be represented in many different ways. The goal of this paper is to design an automated system that could successfully reproduce designers’ aesthetic evaluation of a document page layout. We propose using a graphical interface, where designers can adjust aesthetic parameters of a document layout. Their output is then stored as a feature set and used to fit a probability distribution. This task is not trivial, since designers do not think in terms of a mean and covariance parameters of the parameters they adjust. Allowing them to have a freedom to change many aspects such as relative position and number of different types of document page elements would make our task very difficult. We address this issue by fixing a template and allowing designers to change only a limited number of parameters, and repeat this process for a different document template. Having these constraints it is reasonable to assume that the aesthetics probability distribution can be defined as a Gaussian. Similar approach is used in speech recognition literature, where a speech is represented as a sequence of phonemes or speech segments and each segment is modelled as a Gaussian distribution. [Ostendorf and Roukos 1989]

The authors in [Damera-Venkata et al. 2011] used a concept of a probabilistic page template. An example of such template is illustrated in Fig. 1. The probabilistic template represents only relative positions of page elements, that are dependent on values of probabilistic aesthetics parameters. In description of their algorithm the authors set up prior probabilities on the variables \( \theta_i \), modeling them as independent Gaussian random variables. Therefore, for a particular template \( t \), the precision matrix of the variable vector \( \theta_t = [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5] \) is set to be \( \Lambda_t = diag [\lambda_1, \ldots, \lambda_5] \), where the entries \( \lambda_i \) are assigned deterministically. However, for an accurate overall document page aesthetics assessment the concept of independency of parameters is not correct, and the precision matrix \( \Lambda_t \) should be full and store covariance factors between different parameters \( \theta_i \).

In this paper we propose a method of learning a multivariate Gaus-
sian probability distribution of aesthetics parameters $\theta_t$ from designers’ input. We initiate the learning step by performing adjustments to the layout for various content amounts. First, we choose a template $t$ from a library of all possible document templates. Then, for the given template, we place a document layout illustration with fixed amount of content on a graphical user interface and allow designers to change the aesthetics parameters. The layout on the GUI changes with every adjustment designer makes for possible values of selected parameters. The images in the layout are represented as green areas and the text regions are filled with dummy text. An illustration of a GUI is shown in Fig. 2. If the designer chooses to see what the actual document looks like for the layout s/he has created, s/he has an option of generating a pdf document. Once all the variable parameters are set to values that correspond to the created, s/he has an option of generating a pdf document. After finalizing the template, we proceed with fitting a Gaussian model fit to the whole feature set. In order to obtain the ultimate Gaussian model fit to the whole feature set we begin with fitting Gaussians to the feature vectors of the same quality rank $\theta_j$. After simple mathematical manipulations we come up with the cost function:

$$f(\theta_j) > f(\theta_{j+1})$$

(1)

where

$$\theta_j \in \Theta_j$$

$$\theta_{j+1} \in \Theta_{j+1}$$

$$1 \leq j \leq M - 1$$

To achieve this, we require our Gaussian function be such that the feature vectors of the same class are allocated along the same contour. That is we would like the Gaussian parameters to satisfy:

$$\{\overline{\theta}, \Sigma\} = \arg\min_{\overline{\theta}, \Sigma} \sum_{\theta_i \in \Theta} \frac{1}{N_{\theta_i}} - f(\theta_i)$$

(2)

, where $N$ is the normalization constant of the Gaussian distribution. Note, that in case $r_i = 1$ we obtain $f(\theta_i) \approx \frac{1}{2}$, which means that $\theta_i$ is at the proximity of the mean $\overline{\theta}$. After simple mathematical manipulations we come up with the cost function:

$$\{\overline{\theta}, \Sigma\} = \arg\min_{\overline{\theta}, \Sigma} \sum_{\theta_i \in \Theta} 2\log r_i - (\theta_i - \overline{\theta})^T \Sigma^{-1}(\theta_i - \overline{\theta})$$

(3)

In order to obtain the ultimate Gaussian model fit to the whole feature set we begin with fitting Gaussians to the feature vectors of the same class. The mean and the covariance matrix of a local Gaussian fitted to feature vectors of class $\Theta_j$ can be calculated as ML estimators:

$$\overline{\theta}_j = \frac{1}{|\Theta_j|} \sum_{\theta_i \in \Theta_j} \theta_i$$

(4)
\[
\Sigma_j = \frac{1}{|\Theta_j|} \sum_{\theta_j \in \Theta_j} (\theta_j - \bar{\theta}_j)(\theta_j - \bar{\theta}_j)^T.
\]

where |\Theta_j| denotes the cardinality of a class \Theta_j. We assign a weight \(w_j\) to each of the local Gaussian distribution. Then the overall mean and covariance matrix of the whole feature set can be calculated as:

\[
\bar{\theta} = \left( \sum_{j=1}^M \sum_{\theta_j \in \Theta_j} w_j \Sigma_j^{-1} \theta_j \right) \left( \sum_{j=1}^M w_j \Sigma_j^{-1} |\Theta_j| \right)^{-1}
\]

\[
\Sigma = \sum_{j=1}^M \sum_{\theta_j \in \Theta_j} w_j \Sigma_j^{-1} \theta_j
\]

After plugging equations (6) and (7) into (3) we come up with a linear system

\[
(\theta_1 - \bar{\theta})^T \Sigma_1^{-1} (\theta_1 - \bar{\theta}) \ldots (\theta_1 - \bar{\theta})^T \Sigma_M^{-1} (\theta_1 - \bar{\theta})
\cdot \ldots \cdot
(\theta_L - \bar{\theta})^T \Sigma_1^{-1} (\theta_L - \bar{\theta}) \ldots (\theta_L - \bar{\theta})^T \Sigma_M^{-1} (\theta_L - \bar{\theta})
\cdot \begin{pmatrix} w_1 \\ \vdots \\ w_M \end{pmatrix} = \begin{pmatrix} 2 \log r_1 \\ \vdots \\ 2 \log r_M \end{pmatrix}
\]

(8)

be rather unrealistic to assume that this algorithm provides with the solution that satisfies the condition we set in (1) for all of the feature vectors without exception. Hence, to handle the outliers we implement Random Sample Consensus (RANSAC) algorithm [FISCHLER and BOLLES 1981]. We randomly divide the feature set into two parts. The first part containing hypothetical inliers, is used to train the model. The second part is used to test against the fitted model and, if a feature vector is tested well, we also consider it as a hypothetical inlier. If sufficiently many vectors are classified as potential inliers, we reestimate the model from the hypothetical inliers and store the model parameters. We repeat this procedure a number of iterations, and choose the best model from the stored ones.

### 4 Experimental evaluation

We train our model and compare the output with aesthetics prior model used in [Damera-Venkata et al. 2011]. The intention in experimental evaluation is to demonstrate the superior performance of learned Gaussian parameters compared to the heuristically predetermined parameters which were implemented in [2011]. For this reason we used the same document templates. In order to fully demonstrate the algorithm performance, we selected the templates with different position of landscape and portrait images, spanning one and two columns of a three column document page layout. For each of the templates, we generated 100 samples of aesthetics feature vectors. Each of the feature vectors was given a rank \(r_j\), that ranged from 1 to 5. The feature sets were randomly partitioned into training and testing sets. For each rank \(r_j\), 25% of the corresponding feature vectors were randomly selected and added to the testing set, and the rest were used in the training.

For testing evaluation we calculated the Root Mean Square Error (RMSE) by slightly modifying the cost function (3). In order to visualize the error terms, for each of the feature classes \(\Theta_j\) in the testing set we calculated the deviation between the actual and estimated ranks:

\[
RMSE(\Theta_j) = \sqrt{\sum_{\theta_j \in \Theta_j} \frac{|r_j - \exp((\theta_j - \bar{\theta})^T \Sigma_j^{-1} (\theta_j - \bar{\theta}))/2|^2}{|\Theta_j|}}
\]

Aesthetics parameters in the experiments corresponded to the heights of document images, whitespaces in the bottom of the page and above the lower image. The RMSE for each of the classes of the three template feature testing sets are shown in Table 1.

<table>
<thead>
<tr>
<th>Template 1</th>
<th>Template 2</th>
<th>Template 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Theta_1)</td>
<td>0.3609</td>
<td>0.2404</td>
</tr>
<tr>
<td>(\Theta_2)</td>
<td>0.6892</td>
<td>0.6691</td>
</tr>
<tr>
<td>(\Theta_3)</td>
<td>0.7442</td>
<td>0.6248</td>
</tr>
<tr>
<td>(\Theta_4)</td>
<td>0.6628</td>
<td>0.7213</td>
</tr>
<tr>
<td>(\Theta_5)</td>
<td>1.2401</td>
<td>1.1230</td>
</tr>
</tbody>
</table>

**Table 1:** RMSE values for the feature classes. Each row corresponds to the testing feature set of the templates. Each column corresponds to the classes of ranks 1 through 5. Note that with the definition (10) the RMSE values can be regarded as deviations from the actual integer values of ranks.

An illustration of the comparison between fitted and predetermined models is shown in Figure 3. Note that since correlations of the aesthetic parameters were not considered in [2011], the corresponding Gaussian has contours that are horizontally inclined. On the other hand, the contours of the fitted Gaussian distribution correctly describe allocation of the feature vectors.
5 Conclusion

In this paper we described an algorithm of fitting a Gaussian model to aesthetics parameters of a document layout. During data gathering step, the designers are provided with a graphical user interface, where they can adjust document layout parameters as well as give overall aesthetics ranks to the resulting layouts. The feature set is used in training a Gaussian distribution that can be used in automated layout composition described in [2011]. This model stores correlation of the aesthetics parameters in full covariance matrix, which turns out to be very important in document composition. After repeating the learning step for sufficiently large template library, the model can be successfully used in quantifying aesthetic value of single page documents.

References


