Classifying High Quality Photographs by Creative Exposure Themes

Supheakmungkol SARIN and Wataru KAMEYAMA
Graduate School of Global Information and Telecommunication Studies, Waseda University
mungkol@fuji.waseda.jp, wataru@waseda.jp

Abstract
In this paper, we propose to utilize contextual camera setting parameters at the time of capture to perform the classification task of high quality photographs. With supervised machine learning algorithm, we build a model that can classify high quality photographs into six creative exposure themes which are commonly known and used by the professional photographers. Our experiments give us an encouraging result.

Keywords: Aesthetics, Experimentation, Classification

1. INTRODUCTION
In this age of the digital photograph explosion, media companies - especially stock photo agents, advertising and printing companies - have huge collections of high quality photographs. The task of selecting a suitable picture for a targeted theme is, and will still be, a burden, even though there are annotations in the collection. For instance, how does one select an image that depicts freezing action, an image that has a great depth of field or an image that implies motion for a front cover of a magazine? To lessen this difficulty, we are looking at the problem of classification from the professional photographer’s perspective.

2. CONSIDERATIONS
2.1. Exposure and Patterns in High Quality Photographs
In photography, the exposure control being a process of controlling light projecting to camera’s digital sensor is the main actor to successful photography. Exposure is determined by three settings - shutter speed, lens aperture and ISO. Correct combination of these three will result in a good photo - a well exposed photo. Obviously, there are many of such combinations that can result in a well exposed photo. However, among them only a few can give interesting photographs. In his book entitled Understanding Exposure [1], Peterson distinguishes seven classes of high quality photographs by exposure theme. He calls them creative exposure themes. Furthermore, he discusses the characteristics and the rules that can be used to produce those images. In this study, we focus only on six exposure themes because we have limited number of photos that correspond to the seventh theme in our dataset. The following explains each theme and Figure 1 shows the example images of those themes.

- **Story Telling (ST)**: when we want great depth of field with all objects inside to be neat and clear. It is usually done using wide angle lens and small aperture.
- **Who Cares (WC)**: when the depth of field is not a concern and when subjects are at the same distance from the lens. It is usually done with middle range aperture.
- **Isolation or Single Theme (I)**: when we want to focus on a specific subject. It is usually done with a large aperture open. Usually, the unfocused part is blur.
- **Freeze action (FA)**: when we want to freeze and capture the moment. This is usually done using very fast shutter speed.
- **Imply motion (IM)**: when we want to convey motion to the audiences. This is usually done using very slow shutter speed.
- **Macro or Close-up (M/C)**: when we want the great detail of the subject or just part of it in close proximity. Usually, we want to record the image from 1/10 to 10 times or more of the actual size. The image often lacks of depth of field.
2.2. Camera Setting Parameters

As described earlier, lens aperture, shutter speed, and ISO play important roles in creating a correct exposure for each theme. Fortunately, unlike conventional camera, current modern digital cameras are equipped with many sensors. Many kinds of information are recorded at the same time when a photograph is taken. If we make an analogy of those sensors to our human eyes, this captured information represents the intention of the (professional) photographers. Usually, when taking a photo, photographer has in mind which type of photo he or she is going to make and configure the camera setting accordingly. Specifically, two main things can be extracted: photographer's intent and the condition in which the image is captured. EXIF specification [2], which is universally supported by most of digital cameras, enables these settings. Some of the important parameters which professional photographers usually refer to and which can be found in the EXIF header of the each image file are: Lens Aperture, ISO, Exposure Time/Shutter Speed, Date and Time, Focal Length, Metering Mode, Camera Model, Exposure Program, Maximum Lens Aperture, Exposure Bias, Flash, etc.

3. METHODOLOGY, IMPLEMENTATION AND RESULTS

With the above considerations, there is an obvious relationship between creative exposure themes and some of the camera setting parameters. Thus, in this work, we propose to categorize the photographs into six creative exposure themes and tackle the problem computationally and experimentally using statistical learning approach by applying on the camera setting parameters.

3.1. Dataset and Extracted Features

We use the recent MIR Flickr 25000 test collection [3]. The photos in the collection are selectively taken from Flickr\(^1\) based on their high interestingness rate. As a result the image collection is representative for the domain of original and high quality photography. 75\% of them have the 5 major settings namely, Aperture, Exposure Time, Focal Length, ISO Speed and Flash. We use all of these features in this work. Based on the camera model found in EXIF, we also distinguish Point-and-Shoot cameras with Digital Single Lens Reflection ones. For our study, a subset of the collection (2736 photos) is labeled into the six themes. The labeling process is done manually based on the strong correspondence of the visual expression of each of the photos to the six creative exposure themes. One problem that we faced during the labeling process is that some photos can be attributed to multiple themes. For that we put the photo to the most suitable class.

3.2. Model Building, Evaluation and Results

We divide our dataset into training (2/3) and testing sets (1/3). We carefully create the random splits within each class so that the overall class distribution is preserved as much as possible. With the training set, several machine learning algorithms such as Decision Tree, Forest, SVM and Linear combination were used to train the dataset and create the models automatically. Finally, to evaluate the models, we test them with the testing set. The confusion matrix is computed. We calculate the performance of each established model by the following measures: precision as percentage of positive predictions that are correct, recall/sensitivity as percentage of positive labeled instances that were predicted as positive, specificity as percentage of negative labeled instances that were predicted as negative, and accuracy as percentage of predictions that are correct. Decision Tree which is rather simpler than other models gives the best performance of all. Due to limited space, we show only our best result. Figure 2 depicts our generated model while Table 1 and Table 2 show the performance of the model.

\(^1\)Flickr website: http://www.flickr.com
Even though we used only the EXIF parameters and the camera type in this study, we obtained an encouraging result. We also observed that the model generated by Decision Tree only use 3 parameters namely, F Number, Exposure Time and Camera Type. Also, the generated model corresponds to what describes by Peterson. This raises question whether more features which might demand high computational costs are needed.

4. CONCLUSION

We present a technique to classify high quality photos from professional photographer’s perspective without using any conventional low-level features. Recently, there have been research efforts in using EXIF metadata to classify and annotate photographs. Ku et al. used scene modes for image scene classification [4]. Sigha et al. utilized optical information for the task of photo classification and annotation [5]. However, to the best of our knowledge, this work on classification of high quality photographs by focusing on the creative exposure themes is the first attempt so far.

Though we obtained a reasonable performance using very few features, for our future work, we would like to see how the integration with other type of features could help this task even more with regards to the trade-off of computational costs. For our immediate study, content-based features such as color, texture, shape, and scene description will be integrated. We also would like to perform our experiment on larger dataset with multiple annotators to avoid any bias.

REFERENCES


TABLE 2: Precision, Recall/Sensitivity, Specificity and Accuracy Rate (Let TP : TruePositive, TN : TrueNegative, FP : FalsePositive, FN : FalseNegative)

<table>
<thead>
<tr>
<th></th>
<th>FA</th>
<th>I</th>
<th>IM</th>
<th>M/C</th>
<th>ST</th>
<th>WC</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>0.71</td>
<td>0.57</td>
<td>0.82</td>
<td>0.46</td>
<td>0.68</td>
<td>0.77</td>
<td>0.67</td>
</tr>
<tr>
<td>Recall</td>
<td>0.09</td>
<td>1</td>
<td>0.65</td>
<td>0.063</td>
<td>0.94</td>
<td>0.98</td>
<td>0.62</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.99</td>
<td>0.82</td>
<td>0.98</td>
<td>0.99</td>
<td>0.92</td>
<td>0.89</td>
<td>0.93</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.84</td>
<td>0.86</td>
<td>0.95</td>
<td>0.87</td>
<td>0.92</td>
<td>0.91</td>
<td>0.89</td>
</tr>
</tbody>
</table>