A Data-Driven Approach to Understanding Skill in Photographic Composition

Todd S. Sachs¹, Ramakrishna Kakarala²,*, Shannon L. Castleman², and Deepu Rajan²

¹ Aptina Imaging, San Jose, CA, USA
² Nanyang Technological University, Singapore 639798
ramakrishna@ntu.edu.sg

Abstract. Photography requires not only equipment but also skill to reliably produce aesthetically-pleasing results. It can be argued that, for photography, skill is apparent even without sophisticated equipment. However, no scientific tests have been carried out to confirm that supposition. For that matter, there has been little scientific study on whether skill is apparent, whether it can be discerned by judges in blind tests. We report results of an experiment in which 33 subjects were asked to use identical cameras to photograph each of 7 pre-determined scenes, including a portrait, landscapes, and several man-made objects. Each photograph was then rated in a double-blind manner by 8 judges. Of those judges, 3 are professional photographic experts, and 5 are imaging researchers. The results show that expert judges are able to discern photographic skill to a statistically significant level, but that the enthusiasts, who are more akin to the general public, are not. We also analyse the photos using computer vision methods published in the literature, and find that there is no correlation between human judgements and the previously-published machine learning methods.

1 Introduction

Photography, like cooking, can be carried out by just about anyone with a minimum of equipment. However, in order to reliably produce aesthetically-pleasing results, skill is also required. While certain aspects of photography depend strongly on equipment, such as colour, focus, and exposure, there is at least one aspect that is a matter of skill: composition. Composition is taught in Arts faculties and is the subject of many books and papers, but few have analysed it from a scientific perspective. That may be because of the difficulty in framing scientific or engineering problems in studying composition. However, there is reason to believe that photographic composition is amenable to both scientific and engineering inquiry, since it depends on spatial arrangements of features and objects.

This paper is devoted to studying skill in photographic composition using a data-driven approach. We report results of an experiment when a group of

* Corresponding author.
photographers, ranging in skill from novices to experts, took photographs of a set of predefined scenes using identical point-and-shoot cameras using identical settings (in full “auto” mode). The photographs were then rated numerically in terms of composition strength by an independent group of judges. We show that a subset of the judges, who are themselves professional photographers, are able to discern skill in a statistically-significant manner.

In order to put our work in context, we review relevant research in the field of photograph aesthetics. It is important to distinguish photographic quality from photographic appeal. As Savakis, Etz & Loui [1] point out, photo quality (sharpness, noise level, dynamic range) has been studied scientifically for decades but, in contrast, photo appeal has been studied very little from a scientific perspective. In [1], the authors experimentally determined the attributes that observers feel are important to deciding which pictures deserve emphasis in a photo album, and found that the most important is composition. In particular, Savakis et al found that composition was much more important (by at least a factor of 3) than either colourfulness or sharpness, two traditional measures of image quality. The photos used in [1] are from ordinary consumers, i.e., there was no segregation into those from professional photographers and amateurs. In contrast, Tong et al [2] explore the distinction of skill, and attempt to classify photographs into those taken by professionals and amateurs using computer vision techniques. Their methods rely on features extracted from the images such as sharpness, colourfulness, contrast, and saliency. The classifier that they develop correlates well (coefficient of 0.85) with rankings given by a group of 16 human observers. However, they do not consider composition as an attribute, nor possible equipment differences between professionals and amateurs. Ke, Tang, & Jing [3] also examine the choice of attributes that distinguish between experts and amateurs, and argue that the “bag of low-level features” approach taken by [2] is not as effective as using high level semantic features. Specifically, Ke et al propose that expert photos are distinguished from amateur shots by the attributes of “simplicity”, which they measure by spatial distribution of edges, “colourfulness” measured by color histograms and hue count, “sharpness” measured from the spatial frequency content, and two low-level features measuring contrast and brightness. Ke et al test their classifier on photos obtained from the website dpcchallenge.net, and find that the sharpness attribute is the most discriminative in distinguishing between the top 10% most highly-rated photographs from the bottom 10% in their test set. However, their study does not consider composition as an attribute.

Datta, Joshi, Li, and Wang [4] propose a machine learning approach to rating aesthetic appeal of photographs. Like the previously-mentioned studies, Datta et al use the attributes of colourfulness, sharpness (depth of field), but, in a novel step, include consideration for composition by using the “rule of thirds”[1], texture, and familiarity (measured by similarity to a group of standard images).

---

1 A well-known maxim in photographic composition is that objects should be placed not in the center of the image, but at one-third or two-third the height or width to draw the user’s attention into the scene.