

Assessing Image Quality on Import From Camera

K. Armin Samii

Computer Science Undergraduate
Univ. of Calif. at Santa Cruz
ksamii@ucsc.edu

Uliana Popov

Computer Science Graduate
Univ. of Calif. at Santa Cruz
uliana@soe.ucsc.edu

Allison Carlisle

Biomolecular Engineering Undergraduate
Univ. of Calif. at Santa Cruz
acarlisle@ucsc.edu

James Davis

Computer Science Professor
Univ. of Calif. at Santa Cruz
davis@soe.ucsc.edu

January 31, 2011

Abstract

In this paper, we propose a novel method for ranking images as they are imported from a users camera. We provide the user with the best image from every scene photographed and a ranking of these scenes. Our goal is to filter images which are technically flawed to allow the photographer to focus on the best image from each similar set. By assuming a chronological import, we rank images relative to each other, eliminating the need for no-reference image based techniques. The set chosen by our algorithm matches the top two choices of photographers in our study with astonishing accuracy (97%).

3. Find: When looking for an image to use from a large set, the user will look for an image that best matches personal preferences and qualities most suited for the task at hand.

Our research has focused on the first step. By assessing the criteria which cause a photograph to be rejected uniformly, we have created an automated system which rejects over 94% of the images a human would reject, based on 50 [1]fifty data sets of 50 images tested on 20 users each. Further, we have created a scale for rating each image's quality on a scale of zero to ten. Over 90% of our ratings were within one of the average user ranking, which was tested on the same dataset of 2500 images and 2000 different users.

1 Introduction

We believe there are three general steps a photographer takes. This serves as a heuristic to determine which qualities are relevant to our research.

1. Flag-or-reject: The user sorts through the images being imported and selects the ones most suited for retouching.
2. Retouch: The user will modify the raw files to stylize and enhance them.

2 Related Work

Various work has been done on the second step (retouching)[...refs] and third step (finding a matching image)[1]. These goals are not shared by our work. Various approaches have been proposed to rank any set of images based on aesthetic or technical quality[...refs] but these do not make use of the reference-based algorithms which are possible when working with images directly from a camera im-

port. A prime focus of other works has been to assess both technical and aesthetic qualities[...refs], but aesthetics are largely preference-based and an automated algorithm may not be trusted by a photographer. Therefore, our work assesses an image based on blur, noise, exposure, and the relationship between colors. Past research has quantified these artifacts globally[...refs] and independently (no-reference algorithms) [...refs]. Our approach combines local-feature algorithms [...refs] and reference-based algorithms [...refs] to extract similar foreground content in near-duplicate photographs and compare that foreground data. This reduces the potential error that would arise by comparing an entire image to its near-duplicate (for example, a change in the background may affect the edge-detection-based blur ranking, even if the foreground remains sharp). Similarly, by separating foreground and background content, we can better measure exposure balance and weight noise levels which closer match the human visual system (HVS)[...refs] and assess exposure levels based on foreground-background histogram comparison, similar to[...refs?].

3 Quantifying Image Quality

Each quality assessed will provide a ranking between zero and nine, with larger numbers indicating higher quality. Because an image which ranks low on any part would be considered poor, we have propose an algorithm which penalizes low scores more than it rewards high scores. For example, a perfectly-ranked noise image (i.e. no noise whatsoever) should still fail to a partially noisy image if it is extremely blurred.

$$\sum_{i=1}^n \frac{\log^{-1}(W_i Q_i)}{9n}$$

Where Q_i represents each quality being ranked, and W_i represents the weight of that quality. Weights are assigned as follows: (list here of how much each part is weighted)

3.1 Content Recognition

We obtain a bounding box around the most salient foreground object, in a method similar to [#]. Rather than using the precision that [#] uses, we aim for accuracy, and observed faster and equally accurate results with this method. We achieve high accuracy because of the simplicity, though secondary objects are sometimes ignored or mistaken as foreground. We have found that this does not heavily affect the outcome of our algorithm.

3.2 Near-Duplicate Detection and Similar-Image Clustering

To find near-duplicates, we first look at the time the photo was taken. The algorithm explained in [#] provides a good estimation of whether or not images were taken sequentially. We use this as a weight for further near-duplicate detection: The first step is to perform a fast 9-segment test similar to [#]. If this does not succeed, we perform a second, more complicated test to determine similarity based on the area and histogram distribution of the foreground[#]. We use the bounding box for this calculation. Along with the weight calculated from the timestamp, we can cluster similar images with high accuracy.

3.3 Blur Detection

Using the bounding box, the blur-detection algorithm quantifies the contrast between the edges and background. It ignores the rest of the image. In a method similar to [#], it obtains a value for the sharpness. This algorithm does not perform well when ranking diverse images, due to differences in scale and levels of acceptable background blur, so constraining it to the bounding box increases accuracy. Furthermore, we compare these rankings to the near-duplicates to remove any photograph-specific artifacts which we may not have accounted for. This increases the accuracy of the algorithm with just one photographs, and does even better when there is a sequence of similar images.

3.4 Noise Detection

We use a binary-weight scheme to weigh the noise, in a weak combination of local and global quality assessment. Noise within the bounding box has more negative weight than noise in the background. Our algorithm is similar to [#], which has results good enough to support itself.

3.5 Exposure

Because there is no universally well-performing model known for well-exposed histograms, the content-separated results provide a more accurate way of examining the photograph. As described in [#], we can accurately rate well exposed images bysome cool method....

3.6 Color Harmony

LA DI DA DI DA

4 Results

With a dataset of 4,000 users obtained through Amazon Mechanical Turk, we have found that our work performs very well. BARSKY!![1] and Luo *et. al*[?] obtained 93% accuracy when classifying images into a binary of professional and non-professional. When classifying "non-professional" as ratings below five and "professional" as ratings above five (ratings of exactly five are ignored), our rating system matches a user's with 96% accuracy. When asking a user to choose which images to keep and which to discard, our algorithm correctly discarded 85% of the images, and incorrectly discarded 18%. When limiting both the algorithm and the user to discarding a fixed number of images (ten), we improved these numbers to 92% accuracy with 6% false-negatives. These results rank favorably with [...ref]. Our last test was to rank images independently. Because this was not the goal of our work, we cannot compare to the state of the art[...ref] which received XX% accuracy, but we did manage a no-reference accuracy of within 30% of the user's rating. Further, when we introduced a single set of twenty near-duplicates to a user and asked

them to rank from best to worst, their top three and bottom three matched our results with 97% accuracy.

5 Future Work

Honestly, there's nothing left to be done. We beat this game. Maybe we can combine our work and Barsky and some photobooth folk in India together to make a superawesome combination of the three steps above. That takes some cutting and pasting, then bam, this conference is over. Great success!

References

- [1] Che-Hua Yeh, Yuan-Chen Ho, Brian A. Barsky, and Ming Ouhyoung. Personalized photograph ranking and selection system. In *Proceedings of the international conference on Multimedia*, MM '10, pages 211–220, New York, NY, USA, 2010. ACM.