SELECTIVE REMOVAL OF EXTRANEOUS PHOTOGRAPHS

K. Armin Samii, Allison Carlisle, Uliana Popov, James Davis

[ksamii,acarlisl]@ucsc.edu,[uliana,davis]@soe.ucsc.edu

ABSTRACT

In this paper, we propose a method for selecting the most representative high-quality images from a set of user photographs. To avoid redundancy arising from many similar images, we find all sets of near-duplicates. We then rank images based on three technical qualities: exposure, blur, and color harmony. Finally, we provide an ordering of the images which accounts for both quality and uniqueness. Our goal is not to rank images based on subjective aesthetic qualities, but instead to help a photographer filter out technically flawed photographs and focus on objectively high quality shots.

Index Terms— image processing, quality screening, exposure, meta data, color content

1. INTRODUCTION

The ease of using digital cameras allows for a large number of photographs to be taken at any given event. This increase in photograph quantity can lead to many unwanted low-quality photographs that need to be filtered by the user, which can be time-consuming and repetitive.

To automate this process, an aesthetics ranking technique could be implemented to find the appeal each image has, but this method would discard many images with potential to be fixed through retouching.

Implementing a technical quality ranking technique will keep the photographs suited for retouching, but does not take into account redundancy of similar images, and the user would be left with a set which either includes many duplicates or does not span the entire event.

We thus propose a technique which ranks images based on their technical qualities with a bias for unique images.

To focus our research and determine which image qualities are relevant to technical image quality, we assume three general steps a photographer takes between shooting and using a picture, which we call the Photographer's Process:

1. Remove: Sort through the imported images and remove the ones least suited for retouching.
2. Retouch: Modify the raw files just selected to stylize and enhance them.
3. Retrieve: Select the retouched images most suitable to a given task.

This work focuses on the removal stage and contains both quality assessment and similarity ranking. We implement technical quality assessment using three modules: blur detection, exposure, and color harmony. The blur detection looks at the gradient magnitude of edges and the similarity to a predicted out-of-focus image. Exposure is measured by finding the balance in brightness throughout the image. Color harmony is based on Cohen's model[1]. We implement similarity ranking using global color similarity, foreground color similarity, and timestamps. These modules are combined to obtain a reordering of the set representing how suited they are for removal (the "Importance Order"). A user study has shown 86% accuracy in our absolute ranking of photographs. The similarity ranking was evaluated separately and found to have 78% accuracy. Fig. 4 shows our accuracy in finding the subject and rating quality.

2. RELATED WORK

Previous works have classified images as professional vs. amateur with high accuracy[2][3][4], which is useful for search engines' retrieval of high quality images. Others have focused on personalized *aesthetic* rankings of photographs[5][6]. All of these works focus on finalized images, while we focus on images that are still raw and may be retouched afterward.

Event classification has been explored in order to choose the best images from an entire event[7][8], while we choose the best images from each scene within each event: images with the same subject content rather than the same event context.

Kormann, Dunker, and Paduschek[9] describe a method of automatically rating and ranking images based on image content and time-metadata, but do not report numerical results, only that their results are better than random.

3. QUANTIFYING IMAGE QUALITY

To find the Importance Order of the user's photographs, we use both quality assessment and similarity clustering. Quality is determined by the rankings of the blur, exposure, and color harmony modules. Similarity of images is determined by matching the color content and timestamp of every pair of images. Both modules make use of a foreground region de-

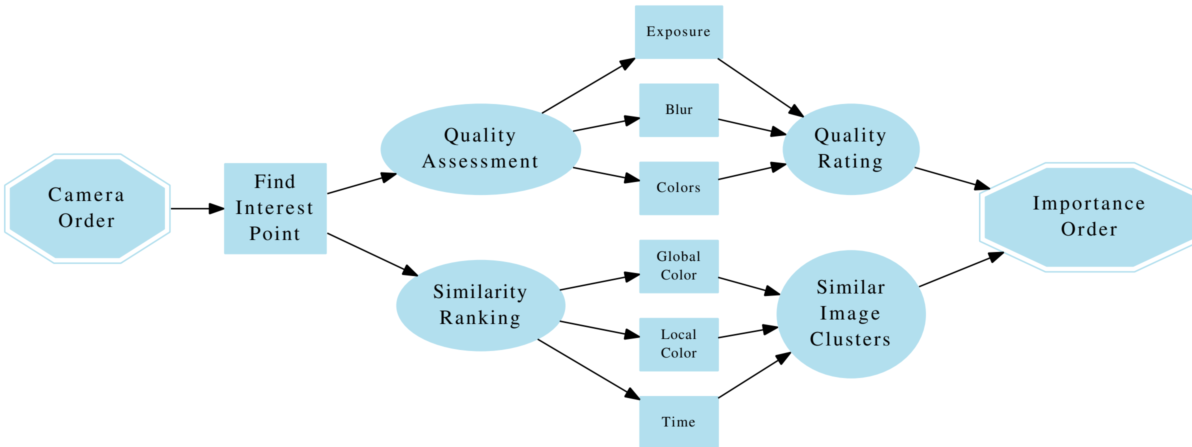


Fig. 1. Interest points are used to find the foreground subject. Blur detection, color harmony, and exposure algorithms calculate an image's quality rating. The color distribution of the local foreground subject, and global image, coupled with the timestamp, cluster images into groups. A reordering of the input results, allowing the user to remove images at the end of this ordering.

tection module based on interest points. We combine these measures to get the final ordering (Fig. 1).

The quality modules (blur, exposure, and color assessment) provide a ranking between zero and nine, with larger numbers indicating higher quality. Because an image which is flawed in any of the three factors would be considered poor, we propose a formula which penalizes low scores more than it rewards high scores:

$$\left(\sum_{i=1}^{n} W_i \left(Q_i + t \right)^{\frac{2}{3}} \right)^{\frac{3}{2}} \tag{1}$$

W_i is weight of module i, n is number of modules, Q_i is module i's rating for the image, and t is a leniency threshold to balance each module's output, empirically chosen to be 4. Weights are assigned as follows: blur 60%, exposure 30%, and color harmony 10%.

3.1. Foreground detection

Several modules make use of foreground detection. We use a Harris Interest Operator (utilizing NASA's Vision Workbench[10]). Interest points are obtained for each image. To extract a bounding box from these points, the most dense rectangle is calculated by maximizing the ratio of interest points to rectangle area. We assume this to be the primary foreground subject.

3.2. Exposure

Exposure is a measure of how appropriate the lighting is in a given image. While easily measured in-camera using a light meter, we must evaluate the lighting condition based on pixel intensity.

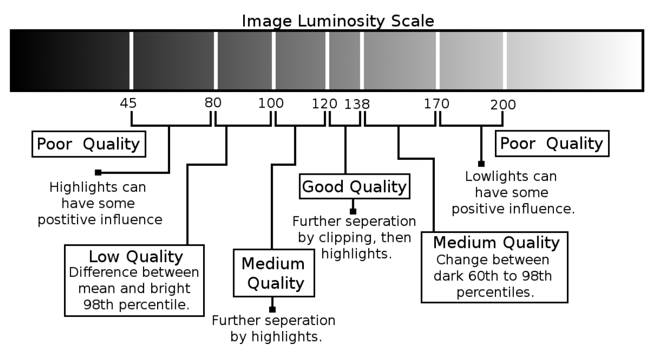


Fig. 2. The exposure process: This figure shows the inital segmentation of images by mean luminosity, then the further segmentation of images of medium and high quality (mean ranges 100-120 and 120-138 respectively).

The measures relevant to exposure quality are: clipping, highlights, lowlights, the upper 60th and 98th percentiles, the lower 60th and 98th percentiles, and variance. Each of these measures is calculated on both the subject and the background of the image, relative to the area over which it is calculated.

Images are first segmented based on their mean brightness value (”mean”). The categories for the mean correspond roughly to a parabolic mapping of exposure values. These divisions, as shown in Fig. 2, are used as the starting point of analysis.

The most extreme mean brightness values are images of very poor quality, and are rated as such based solely on the mean value. Means between 80 and 170 encompass nearly all images of acceptable to excellent exposure, and thus require more analysis. They first are divided into bins according to their mean value, then again divided by the saturation value.

We use the mean value, categories, subcategories, and the eight measures to determine if the exposure is balanced throughout the image. An image with a dark mean should have bright pixels and vice versa. An unbalanced image is rated poorly, even if it has an acceptable mean. Well-balanced images within the acceptable mean range are rated highly.

3.3. Blur Detection

Using the foreground bounding box, the blur detection algorithm first quantifies the spread of the subject's edges. Using the Image Gradient Model proposed in [11], we obtain a value for the blur.

Next, we compute a predicted blurred model of the foreground using a gaussian function. Sharpness is determined by an image's similarity to the predicted blur model, with sharp images being the least similar. These two values combine to determine the amount of blur, which is especially accurate when comparing the foregrounds of two similar images.

3.4. Color Harmony

To determine the quality of a photograph's color content, we use Cohen's seven types of color harmony[1] to determine which, if any, is contained in the photograph. For each image, we find the hue which is most represented across the seven types, as well as the percent of pixels which match any of the harmonies.

We separate the pixels into bins based on their saturation values, using user-study data from Amazon Mechanical Turk to determine the boundaries of each bin. We then look for the most closely matched harmony type within each bin, ranking the bin based on how far it is from the closest type. We combine the bins' rankings for a final quality ranking.

3.5. Similar-Image Clustering

To find similar images we use three measures: time similarity, global color similarity, and foreground color similarity.

First, we use the timestamp to obtain a time similarity index between all pairs of images. Prior work focused on finding large timestamp gaps[12]. We want to allow nonconsecutive images to be grouped together, so we have derived a formula which finds temporal closeness between every pair of images. We find the similarity $S_{i,j}$ between two images i and j on a 0-9 scale:

$$S_{i,j} = \frac{G_{i,j}}{A_{i,j}} \tag{2}$$

Where $A_{i,j}$ is the average gap of the 8 images directly before i and after j, and $G_{i,j}$ is the log of the time gap between the two images taken at time T_i and T_j:

$$G_{i,j} = \log(|T_i - T_j|) \tag{3}$$

The calculated similarity index $S_{i,j}$ is used to weight the results of the next two steps.

Global color similarity is evaluated by dividing the image into a 4x4 grid of average color values. To compare two images, we find the color difference of each square (average difference of the red, green, and blue channels) and deduct from the similarity score if the difference is greater than a threshold of 15%. Because global color similarity is sensitive to subject and camera movement, we then take the bounding box derived in 3.1 and perform the same algorithm on the foreground region, using a threshold of 10%.

Clusters are formed using a Quality-Threshold algorithm. We first put each image into its own group, then iteratively merge the two most similar groups until a minimum similarity threshold is reached. In this way, groups votes for other similar groups, rather than single image voting for other images, resulting in a noise-resistant clustering. This method successfully groups panoramic sets of images together, even when the first image's color does not match the last image's.

4. RESULTS

We gathered Amazon Mechanical Turk user ratings on our own dataset of 459 sequential images.

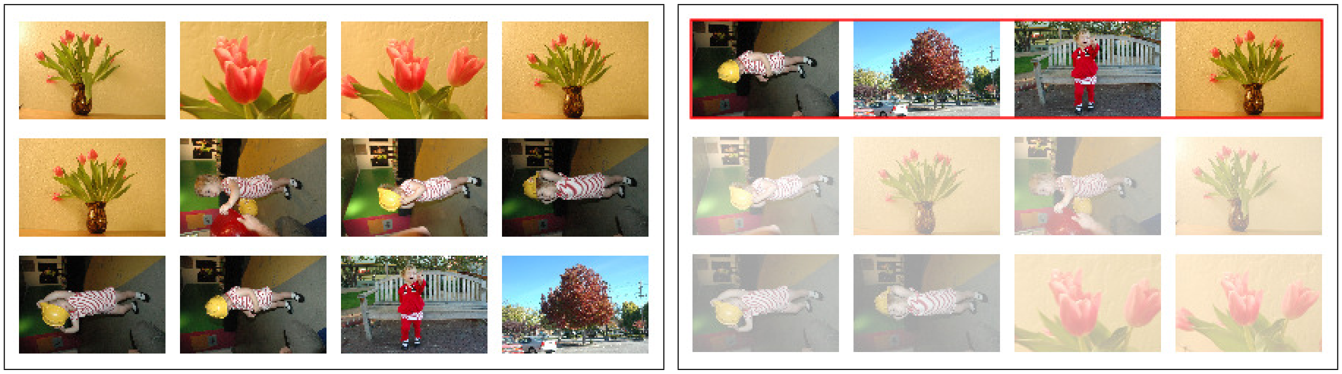
We asked Turk users to group together photographs within four sets of twenty images based on their own personal measure of "similarity." We agree with Turk users' decisions of whether a given pair of images is in the same group with 78% accuracy. We then grouped together photographs with the same subject in focus and achieved an 85% accuracy on the same dataset.

Fig. 3 shows an example of a user's input images being resorted by quality and uniqueness. The top four images in 3b (as would be chosen by a user) represent the highest quality images in each similar-image set. The nontrivial case of zoomed-in flowers was recognized as the same foreground subject. The top images are consistent with Turk users' votes.

5. CONCLUSION AND FUTURE WORK

We propose a method of applying current research to automate another step of the Photographer's Process. By focusing on obtaining an ordering which is representative of all photographs taken, we obtain a diverse set of high quality images similar to what a user would have chosen manually. We derive a novel algorithm for analyzing exposure quality. We improve upon previous algorithms which find temporal gaps between images to obtain a metric for temporal nearness. The final ordering depends on both the quality ranking and the number of similar images which have already appeared.

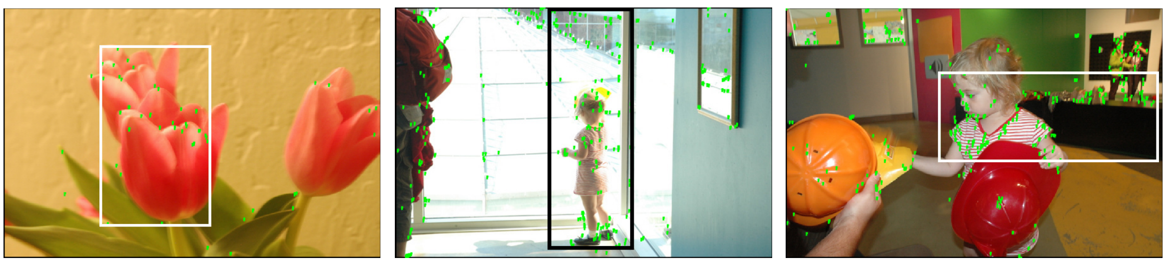
Our work focuses on a small portion of the Photographer's Process. In the future, we would like to see the idea of relative rankings applied to the second Retouching step. (It has already been extensively applied to the third Retrieval step.) Retouching can use relative processing to increase creativity



(a) Unsorted

(b) Sorted

Fig. 3. (a) shows four sets of similar photographs provided by the user. (b) shows the reordered set, which correctly matched the four groups together and chose the top image from that set (as voted by Turk users).



(a) Blur: 2.4 | Exposure: 6 | Color: 4

(b) Blur: 6 | Exposure: 1.2 | Color: 1

(c) Blur: 9 | Exposure: 7.4 | Color: 7

Fig. 4. Examples of (a)(b) low quality and (c) high-quality images. The dots are interest points found; the square is the bounding box considered to be the subject. Despite the too-inclusive box in (c), each algorithm worked properly.

between similar images or combine data from multiple images. With this, we would be able to automate the Photographic Process.

6. PROJECT PAGE

Our code and data set have been made available online at http://www.artoonie.com/imagesorter

7. REFERENCES

[1] D. Cohen-Or, O. Sorkine, R. Gal, T. Leyvand, and Y. Xu, “Color harmonization,” in *Siggraph*. 2006, ACM.

[2] Yan Ke, Xiaoou Tang, and Feng Jing, “The design of high-level features for photo quality assessment,” in *CVPR 2006*, June 2006, vol. 1, pp. 419 – 426.

[3] R. Datta, D. Joshi, J. Li, and J. Wang, “Studying aesthetics in photographic images using a comput. approach,” in *ECCV*, vol. 3953 of *Lec. Notes in Comp. Sci.* Springer Berlin / Heidelberg, 2006.

[4] Y. Luo and X. Tang, “Photo and video quality evaluation: Focusing on the subject,” in *ECCV 2008*, vol. 5304, pp. 386–399. Springer Berlin / Heidelberg, 2008.

[5] X. Sun, H. Yao, R. Ji, and S. Liu, “Photo assessment based on computational visual attention model,” in *Proc. ACM Multimedia*. 2009, MM ’09, ACM.

[6] C.H. Yeh, Y.C. Ho, B.A. Barsky, and M. Ouhyoung, “Personalized photograph ranking and selection system,” in *Proc. of the Int. Conf. on MM*. 2010, MM ’10, ACM.

[7] A.C. Loui and A. Savakis, “Automated event clustering and quality screening of consumer pictures for digital albuming,” *Multimedia, IEEE Transactions on*, vol. 5, no. 3, pp. 390 – 402, Sept. 2003.

[8] Chul-Jin Jang, Ji-Yeon Lee, Jeong-Won Lee, and Hwan-Gue Cho, “Smart management system for digital photographs using temporal and spatial features with exif metadata,” in *Digital Information Management, 2007. ICDIM ’07. 2nd International Conference on*, Oct. 2007, vol. 1, pp. 110 –115.

[9] Daniel Kormann, Peter Dunker, and Ronny Paduschek, “Automatic rating and selection of digital photographs,” in *Semantic Multimedia*, vol. 5887 of *Lecture Notes in Computer Science*, pp. 192–195. Springer Berlin / Heidelberg, 2009, 10.1007/978-3-642-10543-2_23.

[10] M.D. Hancher, M.J. Broxton, and L.J. Edwards, *A User’s Guide to the NASA Vision Workbench.*, NASA Ames, Intelligent Systems Division, 2006.

[11] P. Hsu and B. Chen, “Blurred image detection and classification,” in *Adv. in Mult. Model.*, vol. 4903 of *Lec. Notes in Comp. Sci.* Springer Berlin / Heidelberg, 2008.

[12] J.C. Platt, M. Czerwinski, and B.A. Field, “Phototoc: automatic clustering for browsing personal photographs,” in *ICICS and PCM*, 2003, vol. 1.