SELECTIVE REMOVAL OF EXTRANEOUS PHOTOGRAPHS

Anonymous submission

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 contrast. 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 uses three measures: the gradient magnitude of edges, the similarity to a predicted out-of-focus image, and the probability of a linear point-spread function existing. Exposure is measured by finding the balance in brightness throughout the image, both locally and globally. Contrast is measured based on the local changes in brightness throughout an image. 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 XX% accuracy in our absolute ranking of photographs. The similarity ranking was evaluated separately and found to have 81% accuracy. Fig. 3 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 contrast 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 detection

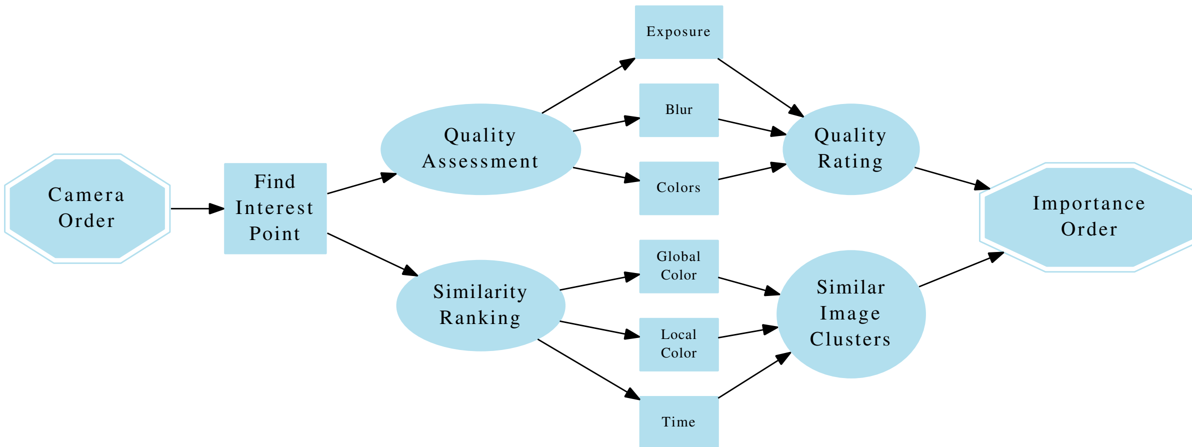


Fig. 1. Interest points are used to find the foreground subject. Blur, contrast, and exposure assessment algorithms are used to calculate an image's quality rating. Images are clustered into groups of similar images based on their histograms, color distribution, and timestamps. A reordering of the input results, allowing the user to remove images at the end of this ordering.

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

The quality modules (blur, exposure, and contrast 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 use a formula similar to the sum-of-squares 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 50%, exposure 40%, and contrast 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.

To assess an image's exposure quality, we use both global and local calculations to compare the average grayscale value of the region to 50% gray (halfway between black and white).

The global method ranks an image highly when the average grayscale value is near 50% gray.

The local method divides the image into 20x20 grid. Images are ranked highly when the bright and dark grid squares average near 50% gray, with no square being extremely right or dark. We also penalize an image if there is not enough variation between grid squares.

The two measures are averaged to provide a final ranking.

3.3. Blur Detection

There are three measures used to assess the amount of blur: edge width, global point-spread, and a comparison against a computed blurry image.

The first two rely on an edge-detection algorithm which is done in three steps. First, we calculate the differences between a pixel's luminosity and the image's average luminosity. Large differences are considered edge pixels. Mislabeled edge pixels are then removed by ensuring each one is connected to another edge pixel. Finally, edges are spaced out by ensuring there is a large radius of non-edge pixels around each set of edge pixels; if there is not, the edge pixel with the largest value from the first step is kept.

The first measure finds the width of each edge by finding a "line of maximum contrast" from the brightest to darkest pixel around each edge. Sharper edges will have shorter line lengths.

The second measure handles looks at the orientation of the lines of maximum contrast. If there are significantly more lines oriented in a single direction than the average, we assume the point-spread to be in that direction.

Finally, if both of these measures are inconclusive, we compute a predicted blurred model of the foreground using a gaussian blur on the image. The closer an image is to this model, the blurrier it is.

These three values combine to determine the amount of blur. Only one needs to have high confidence to assess the blur. The results are averaged if none have high confidence.

3.4. Contrast

UUU

4. SIMILAR-IMAGE CLUSTERING

To find similar images we use three measures: time nearness, histogram similarity, and color distribution.

4.1. Timestamp

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.

4.2. Histogram Similarity

Directly comparing histograms will produce many false positives due to similar-colored but different images, and many false negatives due to changes in exposure affecting the histograms. Our method is robust against both of these.

The image is first divided into a 2x2 grid, with four pairs of values describing each grid square: red, green, blue, and grayscale histograms, paired with the median color for each histogram. The grid protects against false positives. We then divide each histogram into 20 overlapping bins, with each bin containing 8% of the range of values for resistance against minor exposure changes (under 3%).

When comparing the histograms of two images, the bins are scaled so that all channels have the same median value. This provides resillience against larger exposure changes. A similarity ranking for each histogram is derived from the differences in the size of each scaled bin. Finally, each grid square is weighted by the amount of the foreground that it contains. The result is the grid-weighted average of differences between each bin, with smaller differences indicating high similarity:

$$\text{Difference}_{\text{I,J}} = \sum_{g=1}^{4} W_g \sum_{c=1}^{4} \sum_{b=1}^{20} |I_{b,c} - J_{b,c}| \tag{4}$$

Where g is the grid square, W_g is the weight of the square based on foreground overlap, c is the channel, b is the bin, and $|I_{b,c} - J_{b,c}|$ is the difference between the number of pixels in bin b, channel c for images I and J.

4.3. Foreground Color Distribution

We degrade an image to remove features from the foreground while maintaining general color and shape information. First, a strong gaussian blur filter is applied to the foreground to remove features. The image is then scaled to make the comparison have less differences in pixel values. When comparing two images, a smaller pixel-by-pixel difference indicates higher similarity. Because only the foreground is used, this method is robust against frame and subject movement between photographs.

4.4. Clustering

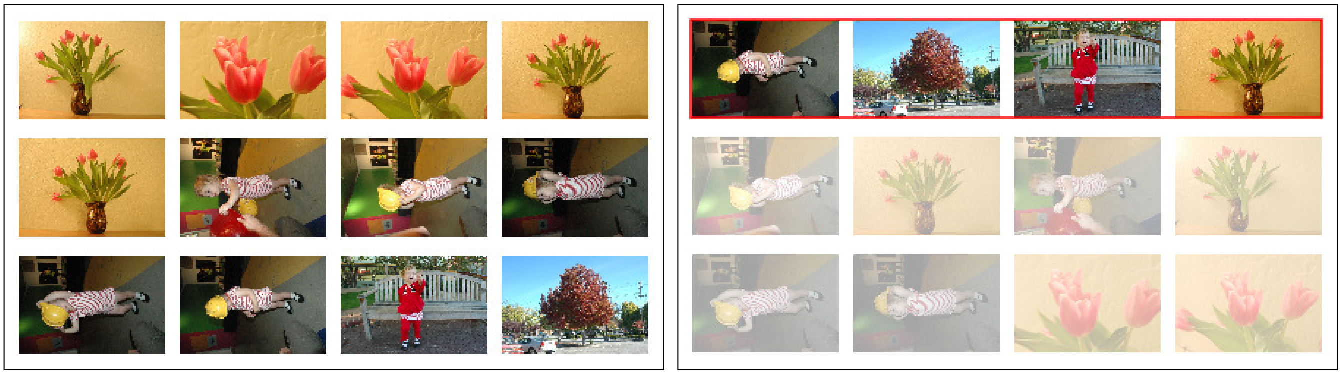
Clusters are then formed using a Quality-Threshold algorithm. We first put each image into its own group, then iteratively merge the two most similar groups until a minimum similarity threshold is reached. In this way, groups votes for other similar groups, rather than single image voting for other images,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

5. 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 81% accuracy.

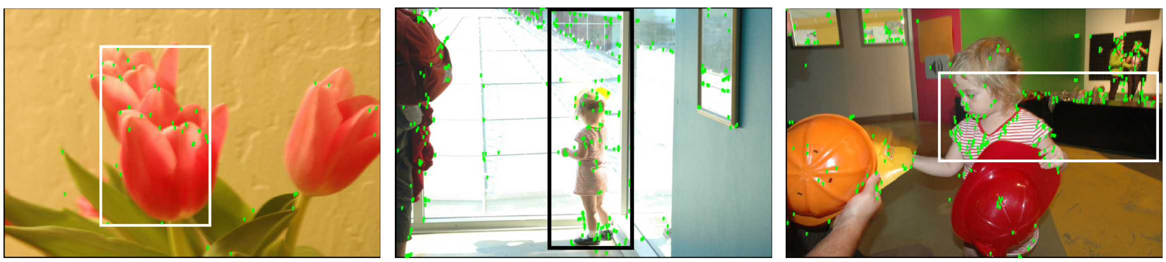
Fig. 2 shows an example of a user's input images being resorted by quality and uniqueness. The top four images in 2b (as would be chosen by a user) represent the highest quality images in each similar-image set. The nontrivial case of zoomed-in flowers was recognized as the same foreground subject. The top images are consistent with Turk users' votes.



(a) Unsorted

(b) Sorted

Fig. 2. (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. 3. 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.

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

7. PROJECT PAGE

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

8. 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.