Automatic Photo Ranking Based on Esthetics Rules of Photography

Wai-Seng Ng[*]   Hong-Cheng Kao[†]   Che-Hua Yeh[‡]   Qi-ya Huang[§]   Wei-cih Jhou[¶]

National Taiwan Unviersity

Bing-Yu Chen[‖]   Han-Wei Shen[**]   Ming Ouhyoung[††]

National Taiwan Unviersity   Ohio State University   National Taiwan Unviersity

Figure 1: *Our algorithm ranks randomly-selected photos from the Flickr website. Here we show two photos (left two) out of the top 100 and three photos (right three) out of the worst 100. A user study of 10 subjects corroborates with our hypothesis that photos can be automatically classified.*

Abstract

Due to the popularity of digital photography, taking, viewing, and preserving photos are much easier. Nevertheless, not many people are familiar with photo esthetics rules, such as composition and color distribution. As an tool, we introduce a quantitative analysis method of photo composition based on well-known photography rules, such as horizon balance, intensity balance, locations of region-of-interests (ROIs), line patterns and merger avoidance. The weighting factors for each of the rules are determined by an experiment involving 500 photos from Flickr sites and dozens of subjects. Support Vector Regression techniques are first used to quantify and "predict" human evaluation, then results from more than 10,000 photos from Flickr are used for the final user study. The user study experiment corroborates with our initial hypothesis that automatic photo ranking is effective. Furthermore, 2000 photos from dpchallenge website are used for training (1000) and test (1000), with two class classification to find the better and worse ones, and we are able to get 80.9% accuracy as compared to human rating.

CR Categories: I.4.3 [Image Processing and Computer Vision]: Enhancement; I.4.8 [Image Processing and Computer Vision]: Scene Analysis

Keywords: Image ranking, photo composition, photography esthetics

1 Introduction

Photography is the process of recording pictures by means of capturing light on a light-sensitive medium, such as films or electronic

---

[*]e-mail: markng@cmlab.csie.ntu.edu.tw

[†]e-mail: photo@cmlab.csie.ntu.edu.tw

[‡]e-mail: chyei@cmlab.csie.ntu.edu.tw

[§]e-mail:

[¶]e-mail:

[‖]e-mail: robin@ntu.edu.tw

[**]e-mail: hwshen at cse.ohio-state.edu

[††]e-mail: ming@csie.ntu.edu.tw

sensors. There are several important issues when taking photos: illumination, exposure, and composition. A photographer determines how he or she wants the subject being illuminated by the environment lighting. A combination of shutter speed, aperture, and ISO sensitivity will be chosen in order to let the film get proper exposed. Before taking any photos, the most important task is to compose the scene: a photographer selects a best view and places different elements into the frame.

Digital cameras become widely available to public in this decade. Since we take so many pictures, it often takes much time to determine which are the better ones to be put in a digital frame or to be printed. We are not always happy with the photos that we took: the balance of the scenery in the photo looks bad or the framing of the scene is not leveled during the shooting, etc. It will be very helpful if we have a tool that can help us determine which photos have better compositions and correct those with bad compositions.

To our knowledge, there are only few well-known researches focusing on photo composition through computational esthetics[Cohen-Or 2008]. Dealing with artistic theories and human visual perception is always a difficult problem, since the results tend to be subjective. Fortunately, photographers around the world have established several general rules for taking good pictures based on their experiences, and lots of them are actually dealing with photo composition. These rules are commonly mentioned in photography literatures [Tsai 1998b; Tsai 1998a; Burian and Caputo 1999; Peterson 2003; Yamaguchi 2006; **?**; **?**] and we select six major rules and convert them into algorithmic functions. In summary, the contributions of our work include:

1. An image analysis framework based on esthetics rules of photography. For each input image, the proposed technique generates a 10-D vector based on six major rules. The score can be regarded as a photo composition descriptor.

2. A validation from two photo websites, dpchallenge and Flickr. In the dpchallenge case, for two class classification, we are able to achieve 80.9% accuracy. In Flickr case, we have conducted a user study, and its result corroborates with the hypothesis that automatic ranking of photos is effective.

2 Related Work

Our technique builds on previous works in re-composition, image information preserving, and automatic image cropping. We will also introduce several important image feature detection tools which will be used in this work.

Re-composition, image information preserving, and automatic image cropping. [Savakis et al. 2000] shows that composition is an important factor to photo appearance. [Banerjee and Evans 2007] have developed an automated composition framework for rules during image acquisition for photographs with one main subject and then it automates photographic composition rules to make the main subject more prominent, such as merger avoidance and reposition. The proposed algorithms are also capable of being implemented on programmable digital signal processors, which presumably can be added into digital still cameras. [Santella et al. 2006] presents an interactive method for cropping photos. A best cropping area is formulated by a general optimization framework with composition rules. The system identifies important contents and compute the best cropping area for any given aspect ratio or size. [Chen et al. 2003] and [Suh et al. 2003] consider preserving the most important visual information under the constraint of limited resolution. [Avidan and Shamir 2007; Wang et al. 2008] presents a technique called seam carving that supports content-aware image resizing functionality for both image reduction and expansion. Google Picasa provide three crop suggestions that are quite impressive, but details are not disclosed[Google Picasa].

Two-class classification. For photographers and home users, how to classify photos into two clusters (good or bad) is a two-class problem[Tong et al. 2004a]. However previous paper only uses several low level features to determine whether a photo is shot by professional photographer or not. Ke et al. [Ke et al. 2006] combined high level and low level features to determine a photo is good or bad, the results are very good, however, mostly the color features are used and didn't include photo composition as high level features. Now, we will focus on the photo composition and the ranking problem in this paper.

Feature detection. The proposed framework employs many individual algorithms to detect edges, lines, region-of-interests (ROIs), and human faces. We use Canny edge detector [Canny 1986] and Hough transform [Duda and Hart 1972] to extract edges and lines in an image. The Canny edge detector is known as an optimal edge detector with several advantages. Hough transform is a well-known line detection algorithm in computer vision literatures. An attention model developed by Itti et al. [Itti et al. 1998] is used to find areas of higher visual attention. In certain types of photographs, such as portraits and group pictures, it is very important to locate where the human faces are. We use the algorithm based on OpenCV, but modified to add color features in the cascade filters to be more accurate (frontal face: 95%, profile face: 80% accurate).

Learning of dataset approach. The paper[Datta et al. 2006] perhaps is one of the papers that are very close to our approach, including using 56 features in photos and use Support Vector Machine in learning. However, being a pioneer in this study, first of all, they don't provide a user study as a formal experiment. Secondly, high level features are either missing or not accurate enough; for example, color harmonization, line patterns, face detection, are not yet used in their paper.

Restrictions of evaluation. First of all, when human faces appear in a photo, they are always the region of interests. We will ignore the attractiveness issue of human faces, although the beauty of a face will definitely affect the user evaluation. Currently we don't try to give human faces *scores* since it is fully discussed in [Leyvand et al. 2008].

In addition to photo composition rules, the solution we employ include color harmonization[Cohen-Or et al. 2006], where the objective function can be used to evaluate the color part of images. Image blur is used as a pre-filter, similarly, photos with under exposure and over exposure are rejected in the first stage, and the rest of the photos are used in the second stage ranking process. Human faces, once detected, are treated as region of interest (ROI), similar to all other objects in a scene, and will not have special scores such as the attractiveness index, since we focus on photo composition only. Finally, we will try to evaluate photos based on photo composition rules and low level features such as contrast and blur, and quantify the parameters in terms of information theory(entropy), and color theory. In total we will use 10 different features (explained in the following sections), including three color/intensity related features (color harmonization, contrast, image blur) and seven photo composition related features. The above features will be classified into six photo composition rules.

3 Rules of Esthetics in Photo Composition

Rules of esthetics in photography describes how to arrange different visual elements inside the image frame. In general, creating a good picture begins with careful placement of basic elements in the scene, together with appropriate lighting and an interesting subject. An experienced photographer will try to cope with visual elements: lines, forms, textures, balance, symmetry, depth, colors, perspective, scale, and lighting. Mastering these rules is the key to produce compelling photos. In this paper we categorize these rules into two major parts: photo composition and color distribution. These rules are further transformed into mathematic functions base on heuristics.

3.1 Photo Composition

There are six rules listed below in photo composition. For each rule, we need to quantify the measurement, and in order to do so, we will introduce information theory and stability analysis first. The above two theories will be used in several rules belows.

Quantification by information theory. In order to quantify our features we are inspired by the information theory used in [Ji and Shen 2006] for dynamic view selection. In information theory, when a word consists of N characters each with the occurrence probability p_i, the average information of this word can be represented as an entropy function in the following,

$$E = -\sum_i p_i \log_2 p_i \qquad (1)$$

(1) Location of ROIs - the rule of thirds. The rule of thirds is one of the most well known rules of photograph composition. Many photographers and artists are aware of the rule of thirds: a photo can be divided into nine equal parts by two equally-spaced horizontal lines and two equally-spaced vertical lines. The four intersection points of these lines and the lines themselves can be used to align important visual elements. This is a very important observation, and we would like to design a weighting function to describe this property. We first use two Gaussian distribution functions which center at $\frac{1}{3}$ and $\frac{2}{3}$ to simulate the influence of rule of thirds. Assume s and t are variables between 0 and 1. The one dimensional weighting function $g'(s)$ is

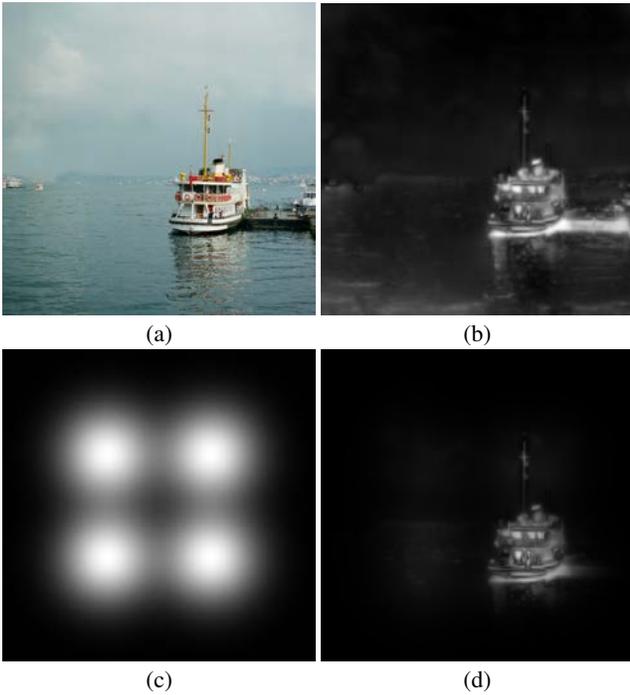$$g'(s) = \varphi_{\frac{1}{3},\sigma^2}(s) + \varphi_{\frac{2}{3},\sigma^2}(s), \qquad (2)$$

Figure 2: *Region of interest.* *(a) The input image. (b) The computed ROI map from (a). Areas of higher intensity contains more information based on the ROI algorithm. (c) The weight map. (d) The combined result.*

where φ is the Gaussian distribution function and σ is the standard deviation. Then the rule of thirds weighting function in two dimension case $g(s,t)$ is

$$g(s,t) = g'(s)g'(t). \tag{3}$$

where $g(s,t)$ is further normalized so that $\max(g(s,t)) = 1$.

We have tried other weighting functions similar to the Gaussian distribution, and the results are very similar.

We mainly use [Itti et al. 1998] method to get the ROI map M_{ROI} of an input image. M_{ROI} is further modulated by the rule of thirds weighting function. The final evaluation value for the rule of thirds Q_{ROI} is

$$Q_{ROI} = \sum_{x=1}^{w} \sum_{y=1}^{h} M_{ROI}(x,y) g\left(\frac{x}{w}, \frac{y}{h}\right). \tag{4}$$

In Figure 2 we show an example of the ROI map.

(2) Horizontal balance. To create a geometrically balanced photo, it is very crucial to keep the horizon balanced. Photos of landscapes, sports, or nautical scenes usually feature absolutely balanced horizons. There are only two tasks in scoring horizon balance: finding the horizon in the image and calculating its tilting angle. We use Canny edge detector [Canny 1986] to create a map that contains all the effective edges. Next, Hough transform [Duda and Hart 1972] is applied to the edge map. The line with the largest response is chosen as the horizon. Then, we find the tilt angle θ_h as the absolute arctangent value of the slope of the horizon. Then the score of horizontal balance is

$$Q_{horizontal} = 1 - \frac{2\theta_h}{\pi}. \tag{5}$$

Figure 3: *Oblique horizon.* *Distant horizon in the left figure is normal while the one on the right is tilted with a skew angle.*

Figure 4: *Line patterns.* *There exists obvious line patterns in two figures, where both parallel lines and radial lines with a possible vanishing point are shown.*

For example, the quality calculated for the better and worse case of Figure 3 are 0.99 and 0.76.

(3) Line patterns. Perspective viewing for man-made artifacts can reveal parallel & radial line patterns, which can help observers to create mental 3D structures. For example, vanishing points in a painting or photo helps maintain 3D illusion, and can extend the depth of view. Further more, line patterns and shapes in photos create special interests. For example, horizontal lines in a photo tends to expand the view.

In our heuristics, photos that have some certain interesting patterns or shapes are often considered good in photo composition. For our evaluation, at first we choose n lines with largest responses from Hough Transform. We then compute the tilt angle θ_t^i of each line. After all, we normalized each angle with the summation of all the angles and we get the probability of each angle:

$$p_i = \frac{\theta_t^i}{\sum_i \theta_t^i}. \tag{6}$$

We then used these probabilities to measure the entropy of the parallel lines:

$$Q_{parallel} = -\sum_{i=1}^{n} p_i \log p_i. \tag{7}$$

The larger value of the entropy, the more parallel lines are in the image. For example, the entropy calculated from Figure 4(a) is 1.0.

The evaluation of the perspective lines is very similar to parallel lines in equation 7. The difference is here we choose the internal angles between lines instead of the tilt angles in the equation 6. We first compute whether these lines are intersected at an vanishing point or not. (see Figure 4(b)). If a vanishing point exists, we compute the internal angle θ_n^i between each line. Finally, we use each θ_n^i to replace θ_t^i and fill them into the equation 6 and compute the $Q_{perspective}$ same as in equation 7. For example, the scores of Figure 4(b) is 0.9.

(a)　　　　　　　(b)

Figure 5: *Area of ROI. The area of bounding box in the left figure is too small, while the right figure area is too big.*

(a)　　　　　　　(b)

Figure 6: *Social Taboo. The head of the person in (b) is separated by a straight line, while the case in (a) is better.*

(4) Size of ROIs. Simpilcity in a photo is shown in [Ke et al. 2006] and it is a most distinguishing factor to determine whether a photo is professional or not. However, the paper [Ke et al. 2006] only used the Laplacian image to compute the size of the area in a photo. Here, we use the size of area of region of interests in a photo, since user always looks at the ROI of this photo and this ROI area have to be affected the user's evaluation. We use the ROI map computed before and project the map onto the x and y axes independently,

$$q_x(i) = \sum_{y=1}^{h} M_{ROI}(i, y), \qquad (8)$$

$$q_y(i) = \sum_{x=1}^{w} M_{ROI}(x, i). \qquad (9)$$

Here, we then find the ROI regions in the q_x and q_y by iteratively discarding boundaries of q_x and q_y so that 95% of the original mass is kept. The area of the bounding box b is used to measure the ROI size score

$$Q_{area} = 1 - \frac{b}{wh}. \qquad (10)$$

For example, the scores of the better and worse case of Figure 5 are 0.98 and 0.01.

(5) Merger avoidance. A merger happens when two or more separate objects merge into one in a photo. Mergers can cause important scene elements losing their own meaning. There are different types of mergers:

1. Geometry: background objects mix or intersect with the subject.

2. Hue: objects with strong colors or similar colors to the subject tends to merge with the main subject.

3. Border: image borders intersect the main subject.

In this section, we only discuss geometric mergers (lines in this case) that are against human subjects in the scene.

We would like to know 1) if there is any human face in the photo, 2) if there is any line in the background which intersects with those face ROIs. First, we use the face detector to look for human faces in a picture. The face detector returns the center c_f and width w_f of a square region that contains a face image if it finds any. Second, we use the same line detection method as described in geometric balance scoring to retrieve lines which contains the largest response. Finally, we check whether the line intersects with any face ROIs or not. Assume the distance between c_f and detected line l_i is d_i, and there are n found lines. The location score s_M is defined as:

$$Q_{face} = -\sum_{i=1}^{n} \sqrt{1 - (\frac{d_i}{w_f})^2}. \qquad (11)$$

3.2 Color and Intensity Distribution

The composition of a photo seems to have strong sense to professional photographer, but the color distribution of a photo is also considered an important part in a photo. In professional photos, the composition and the color distribution of the photos are always combined perfectly. If either one of the two parts is eliminated in one photo, it seems to miss something very important to the users. So we expect the color features to be useful to separate good photos and bad photos. Here, we use three color features: color harmonization, contrast and bluriness.

Color harmonization. We employ the color harmonization technique [Cohen-Or et al. 2006] to measure the quality of color distribution of an image. The defined optimization function is,

$$F(X, (m, \alpha)) = \sum_{p \in X} \|H(p) - E_{T_{m(\alpha)}}(p)\| \cdot S(p). \qquad (12)$$

where H is the hue channel and S is the saturation channel. The input image is X, where each pixel in the image denote p. There are seven harmonic template denote $T_{m(\alpha)}$, where $m \in \{i, I, L, T, V, X, Y\}$ with orientation α. The function $F(X, (m, \alpha))$ is then a function of m and α as in equation 12. Finally, the minimum value of F is the best harmonic scheme of X under the template m and α and this value is used to measure the quality of input image.

For example, the color harmonization index of Figure 7(a) Better and Figure 7(a) Worse is 21313 and 0.

Contrast. A high contrast photo is usually better than a low contrast photo. The contrast is a low level feature and it is useful to evaluate an image quality. Here, we evaluate the quality based on information theory. The entropy of an image is calculated as follows:

1. Convert the RGB image to an grayscale image.

2. Get the histogram from the image and normalize the count of each bin (0 to 255):

$$p_i = \frac{n_i}{n}. \qquad (13)$$

3. Let each bin be a probability value, and all the probabilities can be used to calculate the entropy:

$$Q_{contrast} = -\sum_{i} p_i \log p_i. \qquad (14)$$

The entropy of an image is maximum when the image is balanced in color distribution. That means the color is equally distributed in each intensity. For example, using this algorithm to the better and worse case of Figure 7(b) are 0.99 and 0.3.
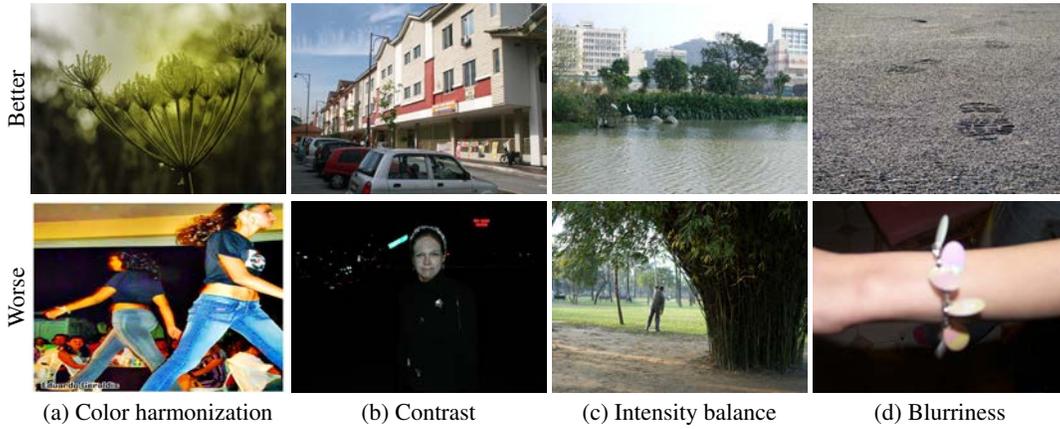
(a) Color harmonization   (b) Contrast   (c) Intensity balance   (d) Blurriness

Figure 7: *Examples of color distribution rules. (a) Color harmonization. The color distribution of the top figure is considered more harmonious than that of the bottom figure. (b) Image Contrast. The intensity contrast of the bottom figure is much lower than the top one. (c) Blur. The bottom figure is blurred for the region of interest while the top one shows a clear focus.*

Intensity balance. Visual elements within a photo can let observers have senses of weight. Keeping the total weight of a photo balanced is important. Large objects generally weigh more than small objects and dark objects weigh more than light colored objects. The position of the elements is also critical. We usually assume the center of a picture corresponds to the center of mass of the picture. Similar to leverage, a heavy weight on one side can be balanced by a lighter weight on the other side if the lighter weight is located at a greater distance from the center.

We can directly use the method from equation 14 to compute the entropy. We calculate the entropy of the left part of image Q_{left} and the right part of the image Q_{right}. And we compare them to the entropy of the whole image Q_{whole}. Now, we can compute the differences between the partial image and the whole image.

$$
\begin{aligned}
d_{left} &= \|Q_{whole} - Q_{left}\| \\
d_{right} &= \|Q_{whole} - Q_{right}\| \\
Q_{balance} &= \frac{1}{d_{left} + d_{right}}
\end{aligned}
$$

If the differences are too large, it means this image is unbalanced. We can see the example from the better and worse case of Figure 7(c), where the scores are 1.51 and 0.82, and so the worse case of Figure 7(c) is unbalanced.

Blurriness. A sharp ROI image is always better than a blurred ROI. For each photo, we employ the blur features proposed in[Tong et al. 2004b; Ke et al. 2006] as a part of our evaluation. First, we convert the input image from spatial domain to frequency domain by Fast Fourier Transform. Each pixel whose power is greater than a threshold is put in the set C ($\theta = 5$),

$$ C = \{(u,v)|\ |F(u,v)| > \theta\}. \tag{15} $$

Finally, count the number of element in C and divide it by the size of the input image. The quality of sharp image is,

$$ Q_{blur} = \frac{\|C\|}{\|I_b\|} \sim \frac{1}{\sigma}. \tag{16} $$

where the σ is the Gaussian smoothing filter parameter. For example, the better and worse case of Figure 7(d) have scores of 0.97 and 0.32.

Subject No.	Percentage of selecting group A photos	Percentage of selecting group B photos
1	0.54	0.46
2	0.47	0.53
3	0.42	0.58
4	0.80	0.20
5	0.64	0.36
6	0.60	0.40
7	0.65	0.35
8	0.51	0.49
9	0.62	0.38
10	0.59	0.41

Table 1: *User study. 10 subjects are used in the evaluation of two groups of photos: A and B, where A contains 100 computer ranked "better" photos, B contains 100 "worse" photos.*

Source	SS	d.o.f	MS	F	p-value
Treatment	0.14112	1	0.14112	12.33	0.0025
Error	0.20608	18	0.01145		
Total	0.3472	19			

Table 2: *Table of ANOVA analysis for Table 1.*

3.3 Accuracy of Individual Features

Although the ranking of photos can be quite difficult because the weighting factors of features need to be determined, however individual feature can be evaluated easily. For example, the following 4 features for photo composition rules such as horizon balance, intensity balance, locations of region-of-interests (ROIs), and merger avoidance, have the detection precision of 71%, 96.8%, 73.1%, 71.4% respectively based on more than 100 sample photos from the Flickr website.

4 Automatic Ranking of Photos

We have two datasets from two famous websites, one from "dpchallenge"(http://www.dpchallenge.com/) and one from Flickr. First of all, we collect higher rated photos and lower rated photos from dpchallenge. This dataset contains about 2000 photos and we split it into training data and test data. The training data have 1000 pho-

tos, half of them are higher rated from dpchallenge user scores. Here, we use [Chang and Lin 2001] library as our main classfication algorithm and use it to test another 1000 photos whether they are good or bad photos. Each photo can be extracted a 10-D descriptor. The SVM library uses the training data which each photos contains 10 features and highest rated photos are labelled as +1 and lowest rated photos are labelled as -1. After the SVM training process with cross validation, the final output accuracy is 80.9671% (787/972) and this result is comparable to Ke et al. [Ke et al. 2006] method (72% accuracy).

A feature selection tool provide by SVM library is used to find which feature is the most effective feature. And we found that using dataset from dpchallenge, bluriness is the most effective feature while evaluating a photo. The second effective feature is perspective line pattern and the third is area of the ROI. Their $F - score$ value are 0.74, 0.027 and 0.025.

Our second dataset is from Flickr which consisted of randomly collect 10,000 photos from the Flickr website. Since there are numerous degraded photos in this data set, we simply filter out bad images which have very low contrast or are deeply blurred in ROI, then we apply our techniques and other features to evaluate all the feature values. Based on support vector regression library[Chang and Lin 2001], four experienced photographers were used in the training phase. Scores from 1 to 5 were given to each photo, and 500 photos from the the source dataset were rated in such a way. Then 10,000 Flickr photos are ranked by our algorithm after support vector regression training. The results are shown in Figure 9 and the recall/precision diagrams are in Figure 8. More detailed data are in our supplemental material.

4.1 User Study

Our training model is computed by support vector regression from four experienced photographers and we use this model to rank our 10,000 photos (without low contrast and deeply blurred in ROI) in an earier version. Here the 10-D vectors are the same, except for "the rule of thirds" function, where another distribution similar to the Gaussian is used. The best 100 and the worst 100 photos from our image set were selected for the second phase user study, which consists of 10 undergraduate students who are randomly chosen. Each experiment consists of two photos on a screen, left and right, from two groups(A,B), and a subject is to determine whether the left or right one is the better looking photo.

Table 1 contains the user study result, where our null hypothesis H_0 is that our algorithm cannot rank photos effectively, and hypothesis H_1 is that automatic classification (better, worse ones) can be effective. Our F-Test value with $p = 0.0025$ and $F = 12.33 > threshold(4.4)$ shows that we can reject our null hypothesis H_0, and the mean of group A is 0.584 while that of group B is 0.416. Therefore, the user study experiment corroborates with our initial hypothesis H_1 that automatic photo ranking is effective.

4.2 Speed Analysis

Total evaluation time for 10,000 photos in generating the 10-D vector scores is 13.8 hours, or 5 seconds/photo, using Intel CoreQuad 2.4GHz CPU PC. Total time of Support Vector Regression analysis for predicting 10,000 photos each with 10-D vectors is 3.5 seconds, or 0.35 millisecond/photo, and it is quite efficient in computation.

5 Conclusions and Future Works

Photo composition perhaps is one of the most important variables that determine whether a photo is good or not. We implement
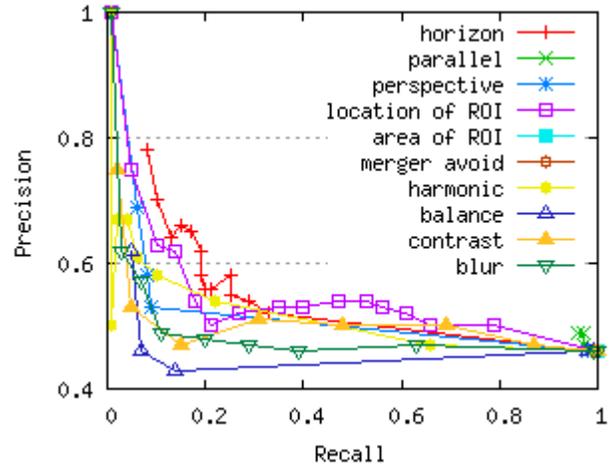
(a)
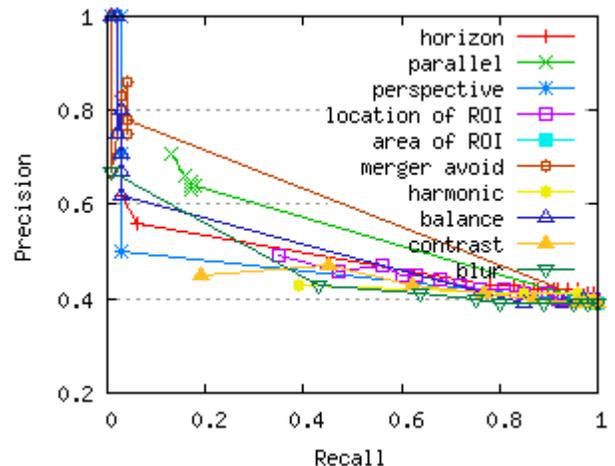
(b)

Figure 8: *(a) The recall/precision diagram of each feature of the good photos. (b) The recall/precision diagram of each feature of the bad photos.*
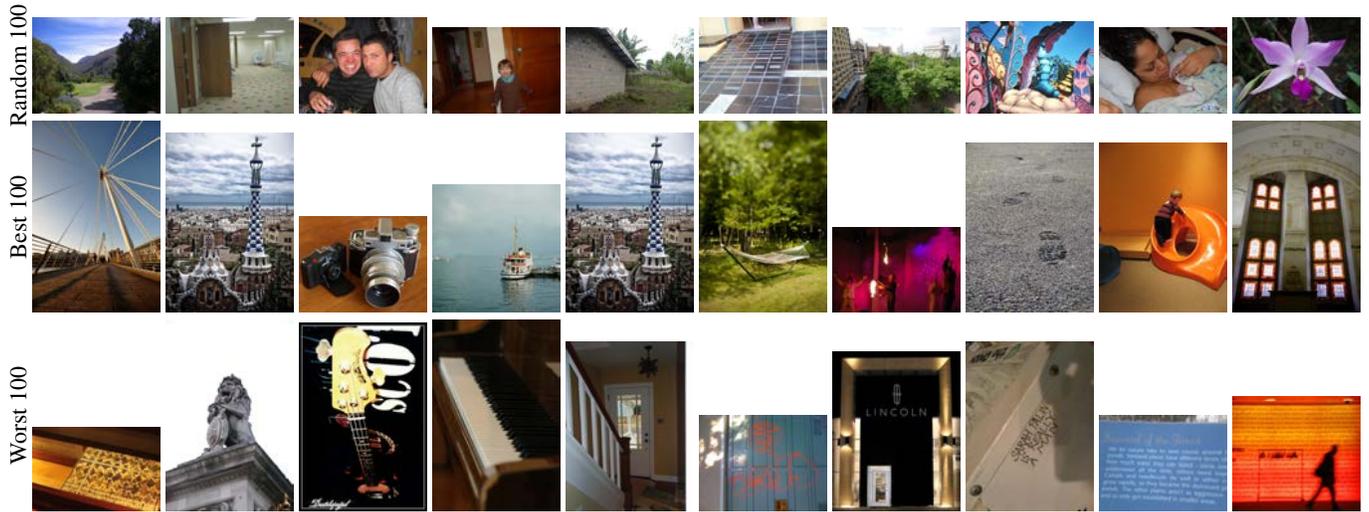
Figure 9: *Photo ranking results. From top to bottom: (1) samples from randomly selected 100 photos from 10,000 photos, (2) the best 100 photos selected by our algorithm from 10,000 photos, and (3) the worst 100 photos resulted from 10,000 photos.*

esthetics rules in photography to automatically analyze six major rules, and use these rules to score photo compositions. Furthermore, according to the analysis results, photos with better composition can be created by automatic cropping. Since we can produce quantitative scores and make recommendations for photos, we hope that an user can gain his/her knowledge during this process and take better photos in the future. Our user study experiment corroborates with our initial hypothesis H_1 that automatic photo ranking is effective.

Although our user study shows that automatic ranking is effective, there exist abnormalities to be discussed. For example, Subject 2 and 3 in Table 1 actually ranked the selected photos in an opposite way as compared to the other eight subjects, which means that our algorithm needs further improvement.

Since photos can be ranked, automatic cropping and rotation is possible by optimization algorithms such as steepest descent algorithm. Preliminary results are given in Figure 10.

References

AVIDAN, S., AND SHAMIR, A. 2007. Seam carving for content-aware image resizing. 10. (SIGGRAPH 2007 Conference Proceedings).

BANERJEE, S., AND EVANS, B. 2007. In-camera automation of photographic composition rules. *IEEE Transactions on Image Processing 16*, 7, 1807–1820.

BURIAN, P., AND CAPUTO, R. 1999. *National Geographic Photography Field Guide*, 2nd ed. National Geographic, 24–33.

CANNY, J. 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence 8*, 6, 679–698.

CHANG, C.-C., AND LIN, C.-J. 2001. LIBSVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

CHEN, L.-Q., XIE, X., FAN, X., MA, W.-Y., ZHANG, H., AND ZHOU, H.-Q. 2003. A visual attention model for adapting images on small displays. *Multimedia Systems 9*, 4, 353–364.

Figure 10: *Automatic correction. For each column, (Top) The original image. (Bottom) The corrected images, where the itemized score improved 0.8 (intensity balance), 0.53 (horizon), 0.16 (location of ROI) respectively, while keeping the rest 9 vectors almost fixed.*

COHEN-OR, D., SORKINE, O., GAL, R., LEYVAND, T., AND XU, Y.-Q. 2006. Color harmonization. 624–630. (SIGGRAPH 2006 Conference Proceedings).

COHEN-OR, D. 2008. Private communication on photo composition.

DATTA, R., JOSHI, D., LI, J., AND WANG, J. Z. 2006. Studying aesthetics in photographic images using a computational approach. In *ECCV (3)*, 288–301.

DUDA, R. O., AND HART, P. E. 1972. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM 15*, 1, 11–15.

GOOGLE PICASA. Three crop suggestions provided by picasa. http://picasa.google.com.

ITTI, L., KOCH, C., AND NIEBUR, E. 1998. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence 20*, 11, 1254–1259.

JI, G., AND SHEN, H.-W. 2006. Dynamic view selection for time-varying volumes. 1109–1116. (Visualization 2006 Conference Proceedings).

KE, Y., TANG, X., AND JING, F. 2006. The design of high-level features for photo quality assessment. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Washington, DC, USA, 419–426.

LEYVAND, T., COHEN-OR, D., DROR, G., AND LISCHINSKI, D. 2008. Data-driven enhancement of facial attractiveness. 1–9. (SIGGRAPH 2008 Conference Proceedings).

PETERSON, B. 2003. *Learning to See Creatively*, revised ed. Amphoto Books, ch. 3, 88–129.

SANTELLA, A., AGRAWALA, M., DECARLO, D., SALESIN, D., AND COHEN, M. 2006. Gaze-based interaction for semi-automatic photo cropping. In *ACM CHI 2006 Conference Proceedings*, 771–780.

SAVAKIS, A. E., ETZ, S. P., AND LOUI, A. C. 2000. Evaluation of image appeal in consumer photography. In *Proceedings of SPIE, Human Vision and Electronic Imaging V*, vol. 3959, 111–120.

SUH, B., LING, H., BEDERSON, B. B., AND JACOBS, D. W. 2003. Automatic thumbnail cropping and its effectiveness. In *Proceedings of the 2003 ACM Symposium on User Interface Software and Technology*, 95–104.

TONG, H., LI, M., ZHANG, H.-J., HE, J., AND ZHANG, C. 2004. Classification of digital photos taken by photographers or home users. In *Proceedings of the 2004 Pacific-Rim Conference on Multimedia*, 198–205.

TONG, H., LI, M., ZHANG, H.-J., AND ZHANG, C. 2004. Blur detection for digital images using wavelet transform. In *Proceedings of the 2004 IEEE International Conference on Multimedia and Expo*, 17–20.

TSAI, M.-F. 1998. *Better Photography*. Recreation Press, 9–15, 17–21, 23–24, 32.

TSAI, M.-F. 1998. *How to Compose a Picture*. Recreation Press, 24–32, 76–88, 108–117, 131.

WANG, Y.-S., TAI, C.-L., SORKINE, O., AND LEE, T.-Y. 2008. Optimized scale-and-stretch for image resizing. 1–8. (SIGGRAPH Asia 2008 Conference Proceedings).

YAMAGUCHI, T. 2006. *Photo Composition Determines All! Methods to Take Great Photos*. Gakken Co., 5–18.