The Design of High-Level Features for Photo Quality Assessment

Yan Ke[1], Xiaoou Tang[2], Feng Jing[2]

[1]School of Computer Science, Carnegie Mellon; [2]Microsoft Research Asia

yke@cmu.edu; {xitang, fengjing}@microsoft.com

Abstract

[1] *We propose a principled method for designing high level features for photo quality assessment. Our resulting system can classify between high quality professional photos and low quality snapshots. Instead of using the bag of low-level features approach, we first determine the perceptual factors that distinguish between professional photos and snapshots. Then, we design high level semantic features to measure the perceptual differences. We test our features on a large and diverse dataset and our system is able to achieve a classification rate of 72% on this difficult task. Since our system is able to achieve a precision of over 90% in low recall scenarios, we show excellent results in a web image search application.*

1. Introduction

We present a top-down approach for constructing high level semantic features for photo quality assessment. Our resulting system can classify between high quality "professional photos" and low quality "snapshots". Informally, we define professional photos as those that would be framed and hung on a wall, and snapshots as those that would stay in a photo album. Having an algorithm for photo quality assessment is useful in many applications. In web image search, the search engine can incorporate a photo's quality into its ranking so that it can return the most relevant and the best looking photos. A home user's photo management software can assess the quality of one's vacation pictures and automatically select the best ones to show. Finally, to see if a computer can perform what has traditionally been a human-only task is an interesting problem in itself.

Unlike previous approaches where a bag of low level features are tossed into a classifier for feature selection, we detail a principled framework for solving this image classification task. A black box approach using low level features give little insight on the reasons why particular features were chosen, or how to design better features for classification. Our approach is to first identify the perceptual
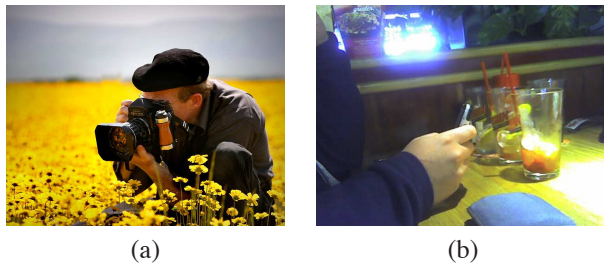

Figure 1. a) "photog in the flowers" by Sara Heinrichs, 2005 and b) "Picture of a picture..." by Ted Johnson, 2005. Many people will agree that the left photo is of higher quality than the right photo. Our system is able to distinguish between these two classes of photos.

criteria that people use for rating photos, and then design features to match the people's perception of photo quality. We then evaluate our system on a large and diverse set of rated photos.

Designing high level features is challenging because even experienced photographers use very abstract terms such as "good composition, color, and lighting," to describe high quality photos. We attempt to distill these abstract concepts into concrete measures, and then design features that computers can use to measure photo quality. Clearly, there is high variation in the perception on the quality of a piece of art. Our work only tries to distinguish between high and low quality photographs, where the classification task is generally easy for humans, as shown in Figure 1. We do not attempt to classify abstract art or photos. We also do not try to discern the finer differences between professional photos, a difficult task even for people. For the remainder of the paper, we will use the terms "high quality" and "professional" photos interchangeably, and similarly with "low quality" and "snapshot".

Digital images sometimes contain metadata that include camera settings such as exposure time, aperture, and ISO. While we believe that such data is useful for predicting image quality [3], we choose to ignore all metadata and only analyze the image content in this work. Only a small fraction of the images currently found online contain metadata, so its usefulness is limited. Our future work will incorporate the metadata where available to improve our results.

2. Related Work

There are a lot of previous works on two-class photo classification. For example, images are classified as photos or graphics [2,7], taken indoors or outdoors [9,14], city or landscape [18], photos or paintings [4], and real or rendered photo [10]. Unlike previous work, where the two classes are well defined, we try to solve a much more difficult problem. Our two classes, high and low quality photos, are subjectively defined with high variations in ratings, and it is not obvious as to what kind of features could be used to differentiate the classes.

Initial work on image quality assessment required the original undistorted image to assess the quality of the degraded image [5,19]. The low quality image is typically degraded by compression or simple noise models. Subsequently, there has been work on directly estimating the quality of a single image [8,13,15]. Unlike our work, these previous work only focused on the quality degradation caused by JPEG compression artifacts. Instead of only looking at the possible low-level distortions that *degrades* an image, we also look at the semantic qualities that make a photo look *good*.

The closest related work is by Tong *et al*. where the authors also try to classify photos as professional or snapshots [17]. There are several limitations in their work. First, because they used the Corel image database, their dataset is fairly homogeneous and therefore it is easy to separate the two classes. Second, they did not analyze the qualities that distinguish professional photos from snapshots and which features are useful for classification. They simply collected a large set of low level features from the image retrieval literature and exhaustively combined them with a standard set of learning algorithms for classification. In fact, they call their own method a "black box" approach. Consequently, their work did not offer insight on intrinsic difficulties of this problem or how to design better features for classification. We will address all of these limitations in our work.

3. What makes a high quality photo?

Before we can design features to assess a photo's quality, we must determine the perceptual criteria that people use to judge photos. We interviewed professional and amateur photographers and non-photographers, and asked them to list, as precisely as possible, the differences between professional photos and snapshots. In addition, we researched photography books [6,12] to see what techniques are often used by professional photographers to raise their quality of work. We found three distinguishing factors between the two types of photos: *simplicity, realism*, and basic photographic technique.

Simplicity. The general consensus among the people we interviewed as to the most distinguishing factor is that professional photos are *simple*. They are simple in that it is obvious what one should be looking at, *i.e.*, it is easy to separate the subject from the background. Snapshots, on the other hand, are often unstructured, busy, and filled with clutter. As illustrated in Figure 1b, it is not obvious what the photo's subject is. There are many ways for professionals to isolate the subject from the background, some of which we discuss below:

- *Background out of focus*. By widening the lens aperture, the photographer is able to blur the background but keep the subject in focus, as in Figure 2a.
- *Color contrast*. The photographer can make the subject pop out by choosing complementary colors for the subject and background, as in Figure 2b. The background typically has very few colors.
- *Lighting contrast*. The photographer can also isolate the subject by increasing the contrast between the subject and the background, as in Figure 2c.

Realism. Another quality that differentiates the two classes is that snapshots look "real" while professional photos look "surreal." Snapshots often depict everyday objects in everyday settings and they capture exactly what the eye sees. Professional photographers, on the other hand, use a wide range of techniques to make their photos atypical and stand out from the snapshots. We list some of the techniques below:

- *Color palette*. Professional photographers are very deliberate in the lighting conditions they choose to photograph under. They might choose to shoot during specific times of the day, *e.g.* morning or dusk, to make the scene look different, as in Figure 3a. Further, they might use colored filters to adjust the color balance, to make the sky bluer, or to make the sunset more brilliant. Finally, the photographers are also careful in the color selection of the scene, for example pairing complementary colors together. Non-photographers taking snapshots, however, will normally take pictures outdoors during the middle of the day, as in Figure 3b, or indoors with a flash. Their scenes will be a mix of random colors and patterns. Therefore, the color palette seen in professional photos and snapshots are likely to be very different.
- *Camera settings*. Most non-photographers use point-and-shoot cameras in the "auto" mode when they take snapshots. The focal length, aperture, and shutter speeds will be in very predictable ranges. Professional photographers will adjust these settings to show different perspectives and create different moods in a photo. For example, a photographer might use a long shutter speed when capturing a waterfall to give it a misty look and a calm feeling, as in Figure 4. While the camera settings can be directly read from the image's metadata, it might be possible to infer the settings when the metadata do not exist.
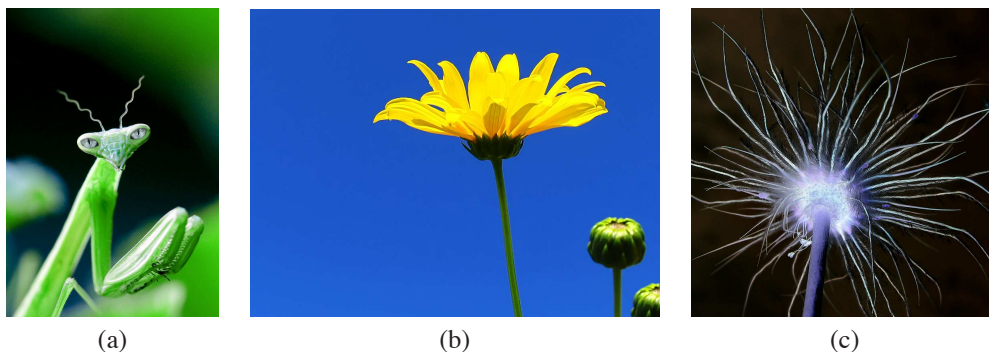
Figure 2. a) "Look Into" by Josh Brown, 2005. b) "Waiting in line!" by Imapix, 2005. c) "alien flower" by Josef F. Stuefer, 2005. Photographers uses many techniques to clearly separate the subject from the background. Some techniques include a) blurring the background, b) using complementary colors (yellow flower against the blue sky), and c) increasing the contrast.
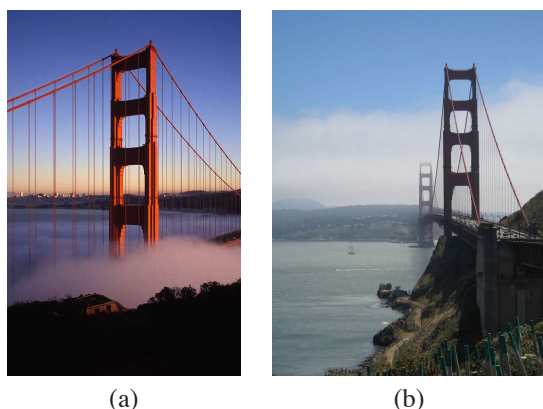

Figure 3. a) "Golden Gate Bridge at Sunset" by Buzz Andersen, 2005 and b) "Golden Gate 3" by Justin Burns, 2005. The lighting conditions during sunset makes the left photo look better than the right photo, which was taken during the day.
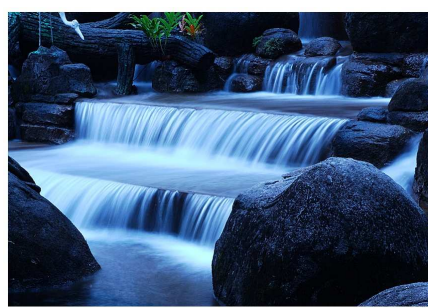

Figure 4. "Somewhere Only We Know Prt2" by Aki Jinn, 2005. The photographer used a long shutter speed to smooth the flow of water. A snapshot using a standard shutter speed would create a very different photo.

- *Subject matter.* An image's subject matter and object pose may be one of the most distinguishing factors between professional photos and snapshots. For example, when professional photographers take pictures of ordinary objects, they usually compose them in unusual poses, or in unusual settings. Although many non-photographers take snapshots of flowers, few will take a picture of a flower from below using the plain blue sky as the background, as shown in Figure 2b. If we can determine that the subject matter or object pose is "unusual", we may be able to conclude that the photo is by a professional. Unfortunately, this is presently infeasible since it requires general object recognition to be solved.

Basic Techniques. There are certain factors that degrade the quality of a photo irrespective of the photo content. We list them below:

- *Blur.* It is extremely rare for an entire photo taken by a professional to be blurry. It is often the result of poor technique, *e.g.* camera shake, or poor equipment, *e.g.*

low quality lens.
- *Contrast.* Professional photos typically have higher contrast than snapshots. Low contrast photos look washed out, and therefore one of their first steps in enhancing a photograph is to increase the contrast so that all 8-bit gray levels between 0 and 255 are used. Non-photographers who use point-and-shoot cameras often end up with low contrast photos due to poor equipment.

4. Proposed Features

Given that we now understand the high level differences between professional photos and snapshots, we propose the following features for classification. Each of the features was carefully chosen to measure some aspect of the criteria listed in the previous section.

4.1. Spatial Distribution of Edges

We compute the spatial distribution of the high frequency edges of an image to try to capture its *simplicity*. Since snapshots often have cluttered backgrounds, we expect the edges to be uniformly distributed in the image. In profes-
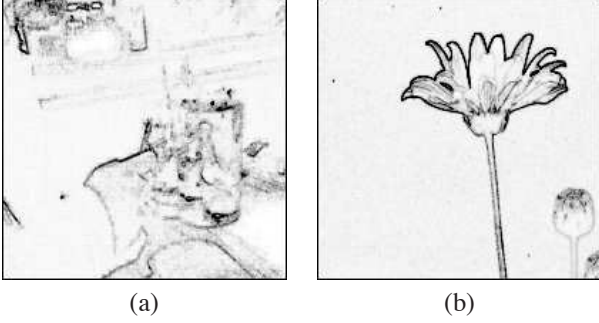
Figure 5. Edge maps from a) the snapshot in Figure 1b and b) the dandelion in Figure 2b. There are more edges near the borders of the snapshot due to background clutter.

sional photos, the subject is well defined, in focus, and may be the only place where high frequency edges are found. Therefore, we expect the edges in professional photos to be clustered near the center of the image, where the subject is usually found.

We implement edge spatial distribution feature extractor as follows. First, we apply a 3×3 Laplacian filter with $\alpha = 0.2$ to the image, and take its absolute value to ignore the direction of the gradients. For color images, we apply the filter to each of the red, green, and blue channels separately and then take the mean across the channels. Finally, we resize Laplacian image size to 100×100 and normalize the image sum to 1. This allows us to easily calculate the edge spatial distribution of the professional photos and snapshots by taking the mean across all the Laplacian images in each set. Let M_p and M_s be the mean Laplacian image of the professional photos and snapshots, respectively. We use the L_1 distance to measure the distance between the probe's Laplacian image, L, and the mean Laplacian images. The quality of the probe image is defined as

$$q_l = d_s - d_p, \text{ where} \tag{1}$$
$$d_s = \sum_{x,y} |L(x,y) - M_s(x,y)| \text{ and} \tag{2}$$
$$d_p = \sum_{x,y} |L(x,y) - M_p(x,y)|. \tag{3}$$

Figure 5a shows the Laplacian image from the snapshot in Figure 1b and similarly, Figure 5b shows the edges in the dandelion photo in Figure 2b. Clearly, there are more edges near the border of the snapshot than the dandelion photo due to background clutter.

Another way to measure the compactness of the spatial distribution of edges is to measure the amount of area that the edges occupy. We calculate the area of the bounding box that encloses the top 96.04% of the edge energy. We expect cluttered backgrounds to produce a large bounding box, and well defined subjects to produce a smaller bounding box. The area of the bounding box is calculated by projecting the Laplacian image L onto the x and y axes independently,

so that

$$P_x(i) = \sum_y L(i,y), \tag{4}$$
$$P_y(j) = \sum_x L(x,j). \tag{5}$$

Let w_x and w_y be the width of 98% mass of the projections P_x and P_y, respectively. The area of the bounding box is $w_x w_y$, and the quality of the image q_a is $1 - w_x w_y$. For Figures 5 a and b, the bounding box area is 0.94 and 0.56, respectively.

4.2. Color Distribution

This feature tries to identify the differences in the color palette used by professional photographers and non-photographers. Each training image is transformed into a high dimensional color histogram, and we use a kNN algorithm to determine whether the probe image is more like a professional photo or a snapshot. We expect that the photos in each set will form non-overlapping clusters, which allows the nearest neighbors of the probe image to determine its class.

We implement the color distribution feature as follows. For each image, we quantize the red, green, and blue channels into 16 values. A $4096 = 16^3$ bin histogram is created which gives the count of each quantized color present in the image. The histogram is normalized to unit length, as the images are of different sizes and thus would give different absolute counts. We use the L_1 metric to calculate the distance between histograms, which we found to give the best result. For each probe image, we find its $k = 5$ nearest neighbors and calculate its quality q_{cd}, where

$$q_{cd} = n_p - n_s, \tag{6}$$

and n_p and n_s are the number of neighbors that are professional photos and snapshots, respectively.

4.3. Hue Count

The hue count of a photo is a measure of its *simplicity*. Most professional photos look more colorful and vibrant than snapshots. Surprisingly, the number of unique hues they contain is actually less than snapshots, although each color may be rich in tones (brightness and saturation levels). A snapshot of a cluttered scene contains many objects, each with its own color. Figure 2 shows three high quality photos with very low hue counts. We calculate the hue count of an image as follows. The hue count for grayscale images is 1. Color images are converted to its HSV representation. We only consider pixels with brightness values in the range $[0.15, 0.95]$ and saturation $s > 0.2$ because the hue calculation would be inaccurate otherwise. A 20-bin histogram

H is computed on the good hue values. Let m be the maximum value of the histogram. Let N be the set of bins with values greater than αm, *i.e.*,

$$N = \{i \mid H(i) > \alpha m\}. \quad (7)$$

The quality of a photo is q_h, where

$$q_h = 20 - ||N||. \quad (8)$$

α controls the noise sensitivity of the hue count and we found $\alpha = 0.05$ to produce good results on our training set. Using this algorithm, we find that the hue count for "photog in the flowers" is 3 while the hue count for "Picture of a picture..." is 11 (see Figure 1).

4.4. Blur

A blurry photo of a scene is almost always worse than a sharp photo of the same scene. In professional photos, there is always some part of the photo that is sharp and in focus. Blur estimation is difficult in itself and there has been previous work in this area [11, 16]. We decided to use both Tong *et al.*'s blur estimation technique [16] and combine it with our own, which we found to give the best results. We model a blurred image I_b as the result of a Gaussian smoothing filter G_σ applied to an otherwise sharp image I_s, *i.e.*,

$$I_b = G_\sigma * I_s. \quad (9)$$

We would like to recover the smoothing parameter σ given only the blurred image I_b. The image quality would be inversely proportional to σ. Let us assume that the frequency distribution for all sharp photos I_s is approximately the same. We can estimate the maximum frequency of the image I_b by taking its two dimensional Fourier transform and counting the number of frequencies whose power is greater than some threshold θ. In other words, let the two dimensional Fourier transform be denoted by

$$F = FFT(I_b). \quad (10)$$

Let the set of frequencies present in I_b be denoted by

$$C = \{(u, v) \mid |F(u, v)| > \theta\}. \quad (11)$$

Since the Gaussian filter G_σ only removes high frequencies, the maximum frequency present in the image is equal to $||C||$. Thus, we define the image quality as

$$q_f = \frac{||C||}{||I_b||} \sim \frac{1}{\sigma}, \quad (12)$$

where the normalizing constant $||I_b||$ is the size of the image. The parameter θ is needed because a Gaussian filter does not produce a sharp cutoff of high frequencies and also for robustness against noise in the image. We use $\theta = 5$ for
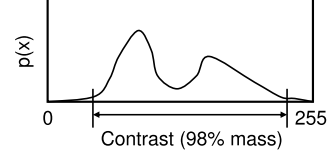

Figure 6. An illustration of how we compute the contrast of an image. We compute its gray level histogram and measure the width of the middle 98% gray level mass.

our experiments. Finally, we combine our blur estimation with Tong *et al.*'s using the classifier discussed in Section 5. The quality calculated for the professional photo and snapshot in Figure 1 is 0.91 and 0.58, respectively.

4.5. Low Level Features

We describe two low level features that are particularly important for photo quality assessment – contrast and brightness. Professional photos usually have higher contrast than snapshots. We measure the contrast of an image as follows. First, we compute the gray level histogram H_r, H_g, and H_b for each of the red, green, and blue channels, respectively. Then, we compute the combined histogram H, where

$$H(i) = H_r(i) + H_b(i) + H_g(i). \quad (13)$$

The combined histogram H, shown in Figure 6, is normalized to unit length, as each image is of different size. The contrast quality, q_{ct}, is equal to the width of the middle 98% mass of the histogram. The contrast calculated for the high and low quality photos in Figure 1 is 254 and 226, respectively.

Most cameras will automatically adjust a photo's brightness approximately 50% gray. However, professionals recognize that the exposure (brightness level) required for the subject and the background could be quite different. Therefore, they will adjust the exposure to be correct on the subject only, causing the average brightness for the entire photo to deviate from 50% gray. In some cases where the background is pure black or white, the deviation could be quite severe. The larger the deviation, the more likely that the photo was taken by a professional. We calculate the photo's average brightness, b, which the classifier can use.

5. Classification

Given our list of quality metrics, q_i, we want a principled method to combine them into an overall quality metric, q_{all}. A naive method might use a weighted linear combination of the quality metrics. However, the quality metrics we defined are not linear, for example a photo's contrast. Therefore, we use a naive Bayes classifier to combine the quality metrics.

Define the overall quality metric,

$$q_{all} = \frac{P(Prof \mid q_1 \ldots q_n)}{P(Snap \mid q_1 \ldots q_n)} \qquad (14)$$

$$= \frac{P(q_1 \ldots q_n \mid Prof)P(Prof)}{P(q_1 \ldots q_n \mid Snap)P(Snap)}, \qquad (15)$$

by Bayes rule. Assuming independence of the quality metrics given the class,

$$q_{all} = \frac{P(q_1 \mid Prof) \ldots P(q_n \mid Prof)P(Prof)}{P(q_1 \mid Snap) \ldots P(q_n \mid Snap)P(Snap)}. \qquad (16)$$

Our data contains equal numbers of professional photos and snapshots, and therefore $P(Prof)$ and $P(Snap)$ can be dropped from the equations. It is important to note that not all of our features are independent, *i.e.* the edge spatial distribution and the edge bounding box area. Further investigation of learning techniques and classifier performance is left for future work.

6. Experimental Results

6.1. Dataset

It is difficult to acquire a large and diverse set of photos for training and testing. Previous work in no-reference quality assessment had very small datasets [13, 15] because manual labeling is labor intensive. Tong *et al.* used a homogeneous stock photo database, the Corel image database [1], for testing [17]. The Corel database has categories, *e.g.* sunset, that are very easy to separate from the rest of the dataset due to its content, and not its quality. Instead, we acquire our data by crawling a photo contest website, DPChallenge.com. It contains a diverse set of high and low quality photos from many different photographers. Further, the photos have been rated by its community of users, giving us ground truth on the quality of each photo. Users can rate a photo from 1 to 10, and we used the photo's *average rating* as the ground truth. From this web site, we crawled a total of 60,000 photos by 40,000 different photographers. Each photo has been rated by at least a hundred users, giving high confidence to the user ratings. The top and bottom 10% of the photos were extracted and assigned as high quality professional photos and low quality snapshots, respectively. Since quality is such a subjective measure, we wanted to use only the photos with clear consensus on their quality and thus we chose to ignore the middle 80% of the photos. From each set, half of the photos ($\approx$ 3000) were used for training and the other half for testing. Some of the photos, especially the high quality ones, contain borders which we removed using a simple color counting algorithm in order to reduce bias in our results.

6.2. Feature Performance

Two of our features, the edge spatial distribution feature and the hue count feature, are non-intuitive and thus
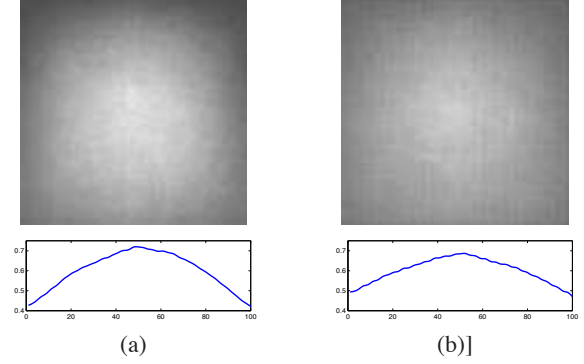

(a) (b)]

Figure 7. Edge spatial distribution for the a) professional photos (M_p) and b) snapshots (M_s). The edges are concentrated more in the center for professional photos and more uniformly distributed in snapshots. There is more background clutter which generate high frequency edges near the borders of snapshots.

we show two results that confirm our hypotheses. Figure 7 shows the edge spatial distribution of the professional photos and snapshots. We see that the distribution is more peaked towards the center for the professional photos, whereas the distribution is more flat for the snapshots. This confirms our belief that the subject is well defined and is usually centered in professional photos. Snapshots, on the other hand, contain a lot of background clutter, which will generate more high frequency edges. When we measured the average hue count of the professional photos and snapshots, we found that the hue count for professional photos is indeed lower. We found that 36% of the professional photos have hue counts of 1 or 2, as opposed to only 27% for the snapshots. This is due to the careful color selection by the professionals, and their desire to keep the photos as simple as possible.

We now give classification results of each individual feature, and also the combined result using our naive Bayes classifier. For each quality metric, we plot a precision-recall curve to show its discriminatory power. We set a threshold on the quality metric and count the number professional photos and snapshots above the threshold. As a reminder,

$$recall = \frac{\text{\# professional photos above threshold}}{\text{total \# professional photos}} \qquad (17)$$

and

$$precision = \frac{\text{\# professional photos above threshold}}{\text{\# photos above threshold}}. \qquad (18)$$

Since we have the same number of professional photos and snapshots in testing set, random selection into the set will give a precision of 0.5. Therefore, the axis showing the precision on our graphs start at 0.5, the worse that any reasonable algorithm can do.

Figure 8 shows the results from each of our quality metrics. We see that all of the metrics do significantly better
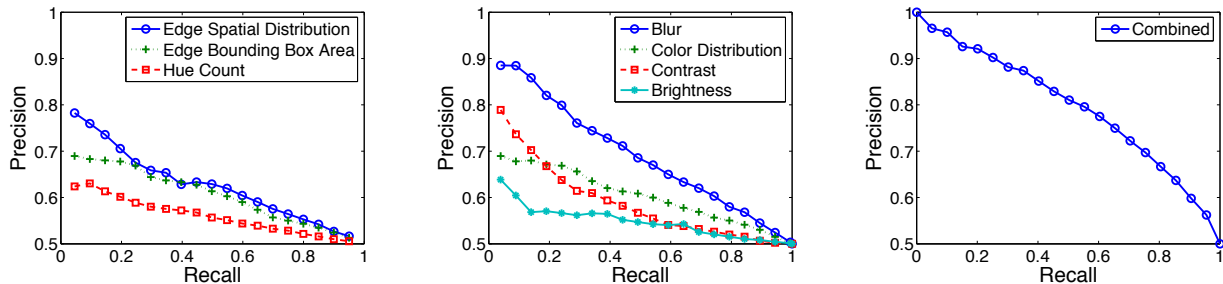
Figure 8. Individual and combined results for each of our quality metrics. The blur feature is the most discriminative metric. The naive Bayes classifier can successfully combine each of the quality metrics to produce the best results, with an equal error rate of 28%. Note that for low recall, our classifier is able to achieve over 90% precision.

than chance in discriminating between professional photos and snapshots. Figure 8 also shows the combined results using our naive Bayes classifier, which is able to achieve an equal error rate of 27.8%. For our dataset, we find that the blur detector is the most discriminative of the quality metrics. This is not surprising since blurry photos are never rated highly, and so if we detect blurry photo, we can be sure that it is low quality.

Using the binaries supplied by Tong *et al*. [17], we compare our methods on the same dataset. Despite using a much smaller set of features, we are able to achieve the same results (27.8% vs. 27.7%). If we combine all of the features and train using Real-AdaBoost, we reduce the error rate to 24.0%. This shows that our features are less complex, just as powerful, and significantly different than previous work.

Because of the subjective nature of this problem, it is inherently difficult to separate the two classes. It is possible for someone to rate a professional photograph poorly, and vice versa. We show the difficulty of this classification task by plotting the ratings distribution of the high and low quality photos in Figure 9. The high and low quality photos were determined by the *average* user ratings, and therefore distribution of the *individual* ratings could overlap. Despite the fact that we only choose the top and bottom 10% of the photos for each of the two classes, there is still significant overlap in the individual rating distributions. However, we expect the class separability to improve if we restrict the testing set to a smaller size, for example the top and bottom 2%. Table 1 shows the results of an experiment where we tested on the top and bottom 2-8% of the images while keeping the training set constant. The error rate dropped to 19% when we test on the top and bottom 2% of the ranked images, which further confirms that our features match the user's perceptual criteria in judging photos.

6.3. Web Image Search

To demonstrate the effectiveness of our algorithm on a real application, we use it to rank the images retrieved from a web search application. For this type of application, the precision of the algorithm is much more important than its
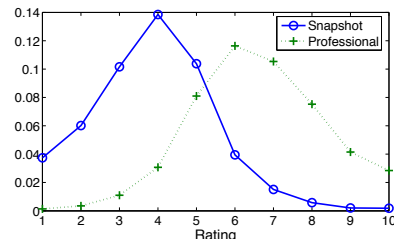

Figure 9. Distribution of people's ratings of professional photos and snapshots. There is significant overlap in the distributions, meaning there is ambiguity in the perceived quality of the photos.

| | Testing on top and bottom n% | | | | |
	10%	8%	6%	4%	2%
Error rate	28%	26%	24%	23%	19%

Table 1. The difference between high and low quality photos are exaggerated when we use a smaller test set. Our error rate decreases as well, which suggests our quality metrics match the perceptual criteria in judging photos.

recall. For example, if a user searches the web for images of German Shepherds, the user only cares that the top 10 images returned are high quality. As shown in Figure 8, our algorithm is able to achieve a precision greater than 90% on the DPChallenge dataset, meaning that 9 out of the top 10 results are high quality photos. Figure 10 shows qualitatively the results of three example search queries. We used Google and Flickr to search for images of "Statue of Liberty", "apple", and "cow". The retrieved images were then ranked by our quality assessment algorithm. We see that *all* top three ranked images are of high quality, and the bottom three ranked images are of low quality.

7. Conclusions

We presented a system for classifying high and low quality photographs. First, we analyzed the perceptual differences between the two classes, and then we designed features to extract those differences. Using a diverse and difficult set of images crawled from the web, we trained our classifier to use our features. On this dataset, we achieve a
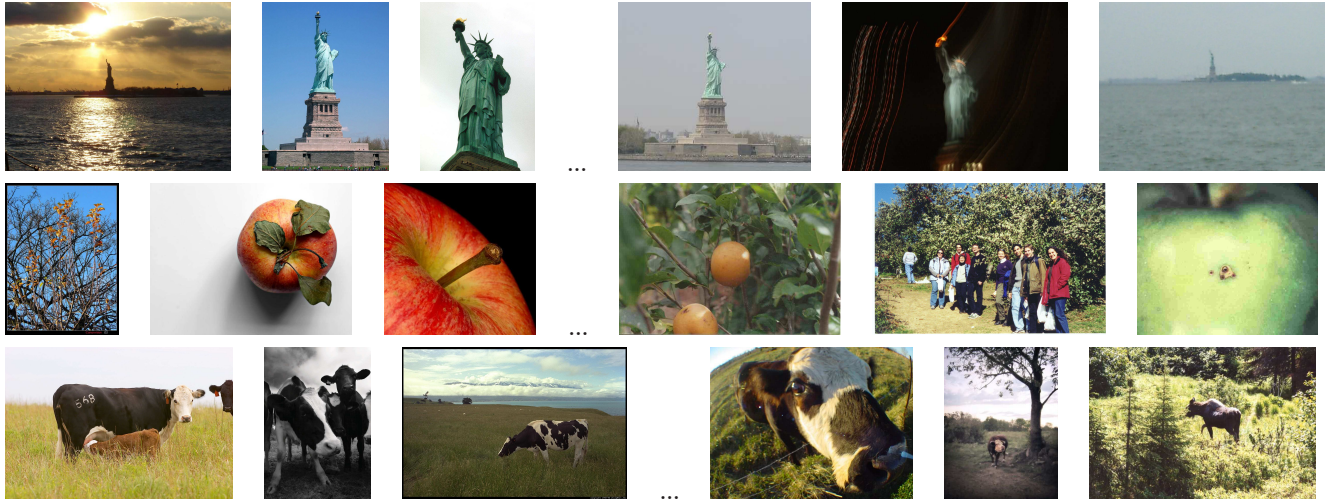
Figure 10. Results from two web image search queries on "Statue of Liberty", "apple", and "cow". The top 3 ranked photos (left) are all of high quality, while the bottom 3 ranked photos (right) are all of low quality.

classification accuracy of 72%.

Our work is not meant to provide a full solution, but rather it aims to inspire more interests in this new and fundamentally important and challenging research direction. Our future work includes designing more features based on the perceptual criteria we discussed, and adding image metadata to give the classifier more information. We also believe that a hierarchical decomposition of the problem would give better results. For example, people portraits and scenery landscapes are very different in content. Therefore, we expect the perceptual criteria used for judging their quality, the photographic techniques used, and consequently the specific features needed to classify them to be all very different. First doing a soft classification of the image's visual category, and then use the category specific features to determine its quality should give better results. Finally, we believe it is important to try to extract high level semantic information from the image. For example, knowing that there is a deep blue sky with wispy clouds in the image should improve the photo's quality rating. Our algorithm can be integrated into existing image retrieval systems as a feature to find high quality photos. Those that include relevance feedback can further fine-tune our algorithm to learn exactly what kind of photos that a user will like.

8. Acknowledgements

References

[1] Corel Images. http://elib.cs.berkeley.edu/photos/corel/.

[2] V. Athitsos, M. J. Swain, and C. Frankel. Distinguishing photographs and graphics on the world wide web. In *CBAIVL*, 1997.

[3] M. Boutell and J. Luo. Bayesian fusion of camera metadata cues in semantic scene classification. In *CVPR*, 2004.

[4] F. Cutzu, R. Hammoud, and A. Leykin. Estimating the photorealism of images: Distinguishing paintings from photographs. In *CVPR*, 2003.

[5] N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and A. C. Bovik. Image quality assessment based on a degradation model. *IEEE Transactions on Image Processing*, 9(4), 2000.

[6] L. Frost. *The A-Z of Creative Photography*. Amphoto Books, 1998.

[7] A. Hartmann and R. Lienhart. Automatic classification of images on the web. In *Storage and Retrieval for Media Databases*, 2002.

[8] X. Li. Blind image quality assessment. In *Proceedings of International Conference on Image Processing*, 2002.

[9] J. Luo and A. Savakis. Indoor vs outdoor classification of consumer photographs using low-level and semantic features. In *Proceedings of International Conference on Image Processing*, 2001.

[10] S. Lyu and H. Farid. How realistic is photorealistic. *IEEE Transactions on Signal Processing*, 37(3), 1999.

[11] G. Pavlovic and A. M. Tekalp. Maximum likelihood parametric blur identification based on a continuous spatial domain model. *IEEE Transactions on Image Processing*, 1(4), 1992.

[12] B. Peterson. *Learning to See Creatively*. Amphoto Books, 2003.

[13] H. R. Sheikh, A. C. Bovik, and L. Cormack. No-reference quality assessment using natural scene statistics: JPEG2000. *IEEE Transactions on Image Processing*, 14(11), 2005.

[14] M. Szummer and R. W. Picard. Indoor-outdoor image classification. In *CBAIVL*, 1998.

[15] H. Tong, M. Li, H. Zhang, J. He, and W. Ma. Learning no-reference quality metric by examples. In *Proceedings of International Conference on Multimedia Modelling*, 2005.

[16] H. Tong, M. Li, H. Zhang, J. He, and C. Zhang. Blur detection for digital images using wavelet transform. In *Proceedings of International Conference on Multimedia and Expo*, 2004.

[17] H. Tong, M. Li, H. Zhang, J. He, and C. Zhang. Classification of digital photos taken by photographers or home users. In *Proceedings of Pacific Rim Conference on Multimedia*, 2004.

[18] A. Vailaya, A. Jain, and H. Zhang. On image classification: City vs. landscape. In *CBAIVL*, 1998.

[19] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment based on a degradation model. *IEEE Transactions on Image Processing*, 9(4), 2000.