



Detection of Out-Of-Focus Digital Photographs

Suk Hwan Lim, Jonathan Yen, Peng Wu
Imaging Systems Laboratory
HP Laboratories Palo Alto
HPL-2005-14
January 20, 2005*

digital
photographs, out-
of-focus,
sharpness, image
analysis

One of the first criteria that a user uses in deciding to use and consume a photograph is whether it is in focus or not. We have developed an effective and efficient algorithm to detect out-of-focus photographs. In this paper, we describe an algorithm that automatically determines if the captured photograph is out-of-focus through image analysis. It uses several global figure-of-merits which are computed from local image statistics. Experimental results show 90% detection rate with 10% false alarms.

DETECTION OF OUT-OF-FOCUS DIGITAL PHOTOGRAPHS

Suk Hwan Lim, Jonathan Yen and Peng Wu

Hewlett-Packard Laboratories

ABSTRACT

One of the first criteria that a user uses in deciding to use and consume a photograph is whether it is in focus or not. We have developed an effective and efficient algorithm to detect out-of-focus photographs. In this paper, we describe an algorithm that automatically determines if the captured photograph is out-of-focus through image analysis. It uses several global figure-of-merits which are computed from local image statistics. Experimental results show 90% detection rate with 10% false alarms.

1. INTRODUCTION

The auto-focus functionality of a digital camera is not always robust. Photographs taken with auto-focus setting can still result in out-of-focus images. For some photos, the entire image is out-of-focus, which may result from the global motion of the camera during capture or from complete failures of auto-focus function (when the digital camera failed to find appropriate focus for any part of the image). In some other situations, only part of the image is focused. For example, the digital camera focuses on the background instead of the foreground if the image is not properly focused. In this case, detecting out-of-focus photos is not a trivial task since it involves determining the foreground and the background. In this paper, we discuss how we tackled the problem of determining if the captured photograph is out-of-focus.

With the advent of digital cameras, taking photographs is becoming a more fun and easier experience. The number of photographs taken each year is growing exponentially partially due to the low cost and easiness of capturing digital photographs. Higher number of captured photographs requires more effort in selecting the photographs for archiving and printing. For example, sorting through tens or even hundreds of photographs taken during a trip to select the photographs to print or save can be a very laborious task. In the selection process, one of the first criteria that the consumer uses to decide to print a digital photograph is whether it is focused or not. Our goal is to develop an automatic algorithm that can detect out-of-focus photographs such that it can at least reduce the number of photographs to be considered.

When the selection process is taken place on a PC, the consumer would display the photographs on a CRT monitor or LCD screen and he/she would have to look at the photograph very carefully and often zoom in to see if the image is indeed focused. It becomes an even more difficult process when the consumer would have to determine the sharpness by looking at the display devices where the spatial resolution is limited. For

example, it is very difficult to judge whether the photograph is focused or not by viewing it on the tiny LCD screen that is attached on a digital camera or a printer. It is very difficult to judge the sharpness of an image when it is spatially downsampled. We believe that automatically detecting whether a photograph is focused or not will be needed in various places in the imaging that goes from capture, management to print/display.

There have been prior researches intended to tackle a similar problem of detecting the sharpness of images. Shaked et al [1] developed an algorithm to measure the overall sharpness of an image to determine how much sharpening should be applied to each image. It estimates the global sharpness of an image, which is provided as a single value per image. However, sharpness of an image may not be uniform throughout the image especially when the depth of focus is small such that some parts of the image are blurry while some other parts are sharp. Thus, this method cannot determine whether the image is properly focused or mistakenly focused on the background. Banerjee et al [2] developed a method to segment the main subject and realize the rule-of-thirds. To segment the foreground, an additional photo with larger aperture is captured and the difference of the frequency content between the two images taken with different apertures is analyzed. A drawback of this method is that it requires an additional image and that it tries to enforce the rule-of-thirds.

In this paper, we extend the sharpness measure estimator described in [1] such that sharpness is estimated locally rather than globally. Extending the method to handle the estimation of local sharpness required optimizing several parameters and making it more robust. In our approach, the image is first partitioned into blocks, and the local sharpness (or the lack of sharpness) of each block is estimated. Then, several figure-of-merits are calculated from the image data and the matrix of local sharpness measures. By analyzing the figure-of-merits, we determine whether the image is well-focused or not. In our current implementation, there are five figure-of-merits which are brightness, color saturation, median sharpness, density of sharp blocks and composition. Note that more figure-of-merits from the capture metadata or via image analysis can be added in order to improve the detection performance of the algorithm.

2. DETECTION OF OUT-OF-FOCUS PHOTOS

The ultimate goal of the out-of-focus detection algorithm is to be able to determine whether the object of interest is focused or not. Sometimes, the digital camera focuses on the background instead of the foreground if the image is not properly focused. In this case, detecting out-of-focus photos is not a trivial task since

it involves determining the foreground and the background. To determine the foreground, we would have to know the intent of the user, which is extremely difficult just by performing image analysis. Thus, instead of trying to solve a very difficult problem of foreground/background detection, we decided to make some assumptions about the foreground and the background, and determine if the captured photo meets our assumptions. The assumptions (prior knowledge) we are currently using are listed as follows:

1. The foreground is always sharp while the background may or may not be sharp.
2. The foreground is likely to be near the center of the image, specifically on the intersections of 1/3 horizontal and vertical lines widely known as "1/3 composition rule".
3. The foreground is typically brighter than the background.
4. The colors of foreground are typically more vivid and saturated than the background.
5. The size of the foreground is not too small

We noticed that the third and the fourth assumptions often do not hold. Exceptions include the blue sky and white snow which occur commonly, so we developed algorithms to detect the blue sky and snow to prevent false alarms. A summary and block diagram of our solution is given below.

1. Divide the image into non-overlapping blocks;
2. For each block, compute local measures;
3. Compute global figure-of-merits from local measures obtained from Step 2;
4. Determine whether the image is well-focused from the global figure-of-merits.

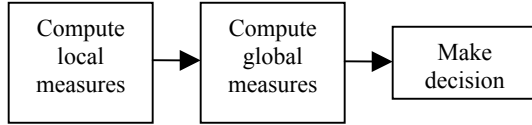


Figure 1: Block diagram of the out-of-focus detection method

2.1 Computing local measures

In Step 1, we currently use the block size of 100 by 100 for the images with the spatial resolution of 2608 by 1952. In Step 2, we first convert RGB values to HSI (hue-saturation-intensity) values and compute local measures such as the sharpness, average brightness and average color saturation. Also, we try to determine whether the block is saturated (because of the sharp reflections and high intensity) and whether the block is part of the blue sky or snow since they create problems for our method.

The sharpness measures of the blocks are computed with a modified version of the method described in [1] using the intensity values. In [1], the key assumption in estimating the sharpness is that natural images are fractals [3],[4],[5] and that the magnitude of the Fourier transform is inversely proportional to the spatial frequency, f , i.e.,

$$|A(f)| \approx \frac{\alpha}{f}$$

,where α is a constant, $a(x)$ is the 1D cross section of the image and $A(f)$ is its Fourier transform. Deviation from this is assumed

to be caused by image blur due to motion blur or focus error from the camera. Using this assumption and Parseval's theorem

$$\int_{f_2}^{f_1} |A(f)|^2 f^2 df = \int_{-\infty}^{\infty} |A(f)W(f)|^2 df = \int_{-\infty}^{\infty} |a'(x) * w(x)|^2 dx = \alpha^2 (f_1 - f_2)$$

,where $a'(x)$ is the spatial derivative of $a(x)$ and $w(x)$ is the ideal band pass filter that only passes frequency between f_2 and f_1 . For $a(x)$, the 1D cross section of natural images, the signal energy after band pass filtering the derivative (or gradient) of $a(x)$ is only proportional to the bandwidth of the filter. Thus, computing the ratio between the high-pass and the low-pass energy of the spatial derivative ($a'(x)$) should only depend on the bandwidth of the filters for ideal images that meet the fractal assumption. For blurry blocks, however, the magnitude of the Fourier transform would drop faster than that of the natural images (fractals), so the ratio of high-pass energy to low-pass energy would be lower than that in a sharp block. As for the implementation, the computation is performed in spatial domain, applying 1D filters along the horizontal pixel lines and vertical columns. Note that the sharpness measure which is the ratio between the high-pass and the low-pass energy of the spatial derivative can be obtained for each line or column. Assuming that the sharpness is uniform within the block, the sharpness value for the block is obtained by averaging the sharpness values of all the lines and columns in the block. Also, since sharpness can only be estimated when there is enough texture (e.g. edges and corners), the ratio is computed only when there is an edge whose spatial derivative is higher than a certain threshold. Note that the computational complexity can be reduced significantly by performing the computation on a smaller set of lines and columns. Since the algorithm requires just filtering and computing the energy, it is very efficient.

In addition, since sharp reflection and flare can cause problems for the local sharpness estimation algorithm, we developed another algorithm that attempts to detect them. Thus, if it was determined that the block has any strong reflectance, the sharpness value for the block was ignored. To increase the robustness of the sharpness estimator, confidence measures were obtained by computing the variance of sharpness values of the lines and columns in the block. In the implementation, sharpness measures that have large variance of the sharpness values within the block were ignored. Figure 2 illustrates the result of running the local sharpness estimator on an image. In the figure, the brightness of the periphery of each block is proportional to the sharpness of each block. Note that many blocks do not have valid sharpness values (shown by the absence of grid in the figure) due to lack of texture in that block.

We also compute average hue, intensity and saturation values for each block. Average intensity value is computed for each block to obtain overall brightness of the block. It is aimed at assessing the third assumption stated at the beginning of this section. Average color saturation is also computed which is to assess the validity of the fourth assumption. Average hue values are computed to detect clear blue sky. In the current implementation, when more than 70% of the pixels in the block have hue values greater than 3.1 and less than 3.8, the block is considered to be part of clear blue sky. The reason clear blue sky is detected is because it is one of the most common exceptions to assumptions 3 and 4.



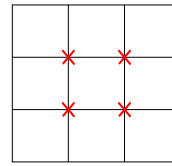
Figure 2: An example of local sharpness estimation result

2.2 Computing global figure of merits

Once the local measures are computed, the global figure-of-merits (Step 3) are obtained. These values try to assess how valid our assumptions are for each image. Since some parts of the image could be blurry while some other parts are sharp, a single sharpness value may not truly represent how well the whole image is focused. Having more than one metric allows us to detect images that are focused on the background in addition to the images that are completely blurry. In our current implementation, we compute 5 figure-of-merits and detect if the image is out-of-focus. The summary of the five figure-of-merits are given here and the detailed description are given in the subsequent paragraphs. “Composition” figure-of-merit is implemented by weighting the local sharpness measures with a function and summing them. It tries to assess the validity of the assumptions 1 and 2. “Brightness index” is obtained by calculating the difference between the average values of the sharp and blurry areas. It tries to assess the validity of the assumptions 1 and 3. “Color saturation index” is calculated similarly for color saturation and assesses the assumptions 1 and 4. “Density of sharp blocks” assesses the assumptions 1 and 5, while “Median of sharpness values” assesses the assumption 1.

First, the Composition (spatial distribution of the sharp blocks) is analyzed. Since it is more likely to have the foreground (or the object of interest) near the center of the image, the figure-of-merit based on the spatial distribution of the sharp blocks was designed to output high scores when the sharp blocks are located near the center. We also wanted the figure-of-merit to include a well-known “1/3 composition rule”, which states that it is good to put the object of interest on the intersections of 1/3 horizontal and vertical lines (See Figure 3). To implement this, we weighted the matrix of sharpness measures with the 2D curve shown in Figure 3 and summed the resulting values to obtain the figure-of-merit which conveys the composition of the photo. This figure-of-merit is crucial in determining whether the image is focused on the background or foreground.

Second, the Brightness index (relative brightness of the sharp blocks) is also analyzed. Since it is more likely for the foreground to be brighter than the background, this figure-of-merit should be high for cases when the sharp areas are brighter than blurry areas. To implement this, we compute the average brightness difference between the sharp and blurry areas.



X: Good Focus spot

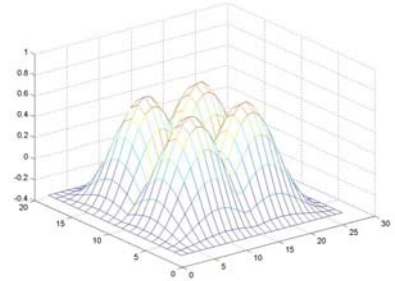


Figure 3: Weighting function for composition

Third, the Color saturation index (relative color saturation of the sharp blocks) is also analyzed. Since it is likely for the region of interest to have more vivid colors than the background, this figure-of-merit should be high for cases when the sharp areas are more vivid in color than the blurry areas. To implement this, we compute the color saturation and subtracted the average color saturation of the blurry areas from that of the sharp areas.

Fourth, the Density of sharp blocks is an important parameter. Here, the density is defined as the number of sharp blocks over the total number of blocks in the image. If the density is too low (<10%), it would mean that only very small part of the image is sharp and that the image is not blurry. If the density is very high (>60%), it would mean that the image is well-focused on both the background and foreground. Note that this measure alone will not be able to determine if the image is focused on the foreground or the background.

Fifth, the Median of the sharpness values of the blocks can show how well the image is focused in general. If this number is high, then the image is sharp overall. We can mark the images as being blurry if this median value is too low. A low value of median could tell if the image has some motion blur caused by the camera or a large object.

Although we did not elaborate, many other statistics measures can be used to determine if the image is blurry or not. We are in the process of enriching the figure-of-merits to improve the performance of our detection algorithms. Also, the use of capture metadata such as exposure time, flash fired and focus distance will be explored more thoroughly in future.

2.3 Decision logic

Once the figure-of-merits are obtained, they are analyzed to determine if the image is well-focused or not. In our current implementation, we have not used a sophisticated training method to optimize the decision rules based on the figure-of-merits, but we plan to use well known data mining and clustering techniques to improve the performance of our method. An example of the decision logic is illustrated with a flow chart in Figure 4. More sophisticated approaches where the figure-of-merits are weighted differently according to confidence measures and sometimes combined together are being under investigation.

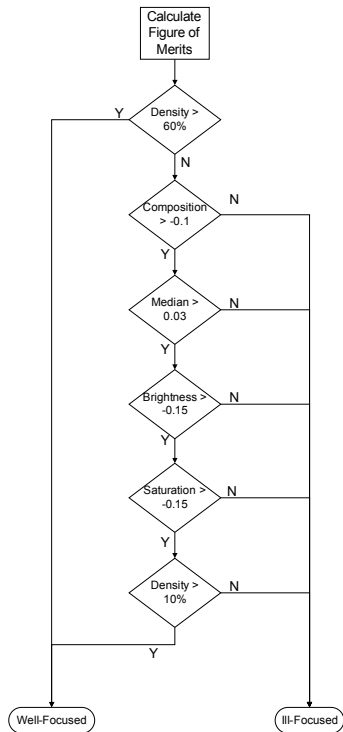


Figure 4: Flow chart of an example decision logic

3. EXPERIMENTAL RESULTS

A human expert laboriously examined each image and determined whether the image is well-focused or not. We used this ground truth and compared it with our simulation results to see how close our algorithm can get to human intelligence. We ran our algorithm on our database of 3000 images (350 of which are out-of-focus). In our implementation with simple predicates, our algorithm can detect 90% of the truly out-of-focus images while producing 10% false alarm. (False alarm means that the photograph is well-focused but our algorithm determined the image is out-of-focus). By using simple predicates (e.g. as shown in Figure 4), we can avoid over-training and keep our algorithm more general.

The misses and false alarms are being analyzed to improve the performance of the algorithm. Misses commonly occur when the sharpness estimator is fooled by direct light sources, sharp reflections or shadows. Extremely small objects of interest and image doubling due to jerky movement of the camera also cause misses as well. (See Figure 5)

False alarms are often caused by objects or sceneries with little texture. Furry animals such as dogs, cats or horses also cause false alarms. In addition, close-up shots of human or smooth objects cause false alarms as well. (See Figure 6)

4. DISCUSSION AND FUTURE WORK

It is worthwhile to point out that the notion of sharpness can depend on the situation and who is looking at the photo. We realized it is important to have flexibility in the algorithm such that it can be used in many use models. We extended the current

method to output how well-focused an image is in multi-levels (e.g. 1-very bad, 5-very good). Another obvious direction is on reducing the rate of misses and false alarms. We are currently using simple predicates to determine out-of-focus images from the five figure-of-merits but plan to use dedicated training tools to come up with more sophisticated and accurate predicates. Another plan is to add more assumptions and figure-of-merits to improve accuracy and robustness. We are also in the process of trying to improve the accuracy by including the capture metadata that are stored during image capture. For example, shutter speed and aperture value may provide us with some clues as to how well-focused our images are.

5. ACKNOWLEDGEMENTS

We wish to thank Doron Shaked for his advice and help on sharpness estimator. We would also like to thank Susan Manson and Murray Craig for providing us with many test photos.

6. REFERENCES

- [1] Doron Shaked and Ingeborg Tastl, "Sharpness Measure: Towards Automatic Image Enhancement", Hewlett-Packard Laboratories Technical Report HPL 2004-84, May 2004
- [2] Serene Banerjee and Brian L. Evans, "Unsupervised Automation of Photographic Composition Rules in Digital Still Cameras", Electronic Imaging Conference (SPIE), January 2004
- [3] Erik Reinhard, "Statistical Approaches to Image and Scene Manipulation", Perceptually Adaptive Graphics: ACM SIGGRAPH/EUROGRAPHICS Campfire, Snowbird, UT, May 2001
- [4] P. Flanagan, P. Cavanagh, and O. E. Favreau, "The human visual system is optimized for processing the spatial information in natural visual images", Current Biology, 10(1), pp. 35-38, January 2000
- [5] Stephane Rainville and Frederick Kingdom, "Spatial-scale contribution to the detection of mirror symmetry in fractal noise", Journal of Optical Society of America A, 16(9), pp. 2112-2123, September 1999



Figure 5: Example of misses

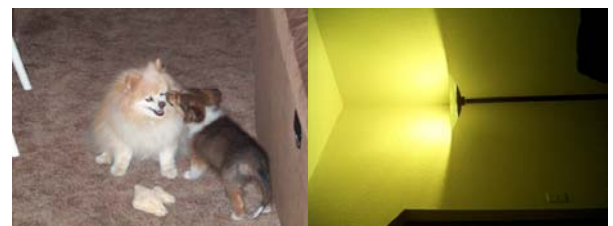


Figure 6: Example of false alarms