Establishment of an application for photographs’ quality detection based on sharpness and light intensity

Hugeng Hugeng1* and Ardis Jevon Tjahyadi2
1Electrical Engineering Department, Universitas Tarumanagara, Jl. Letjen S. Parman No. 1, Grogol, Jakarta 11440, Indonesia.
2Computer Engineering Department, Universitas Multimedia Nusantara, Jl. Scientia Boulevard, Gading Serpong, Tangerang 15811, Indonesia
*hugeng@ft.unutar.ac.id; *Corresponding author’s email

Abstract. Along with the booming of digital photography, camera users grow significantly, especially the users of digital single-lens reflex (DSLR) camera. There are many new users who apply their DSLR cameras for the first time and have little knowledge how to use them. That’s why an application is needed to guide a novice user so that he/she could get better understanding about photography. We have an idea to create such a guidance application that can provide evaluation and suggestion about the quality of user’s photographs. Thus, users can easily develop the ability of photography. Overall, the level of accuracy of this application in identifying deficiencies in the pictures taken is about 75%. This level of accuracy is calculated from errors that occur in detecting both sharpness of the pictures and light intensity (exposure) of the pictures, or any one of them. The errors in detecting the conditions of pictures, occur in some pictures. Hence, it needs further development of methods to improve the accuracy in all conditions of pictures taken. Best accuracy is necessary so that the proper suggestion is given by the application in accordance with the actual condition of the picture.

1. Introduction
Digital photography has made major changes in the field of photography. At first it takes quite a difficult exercise and requires much cost in analogue photography. Besides the picture printing process takes a lot of time to get results from pictures taken. The world of photography nowadays turns into something that easy, attractive and affordable enough to be learned. Photography communities are growing very fast because of the existence of digital photography. Many people with different ages, genders, professions, like to learn photography as their hobby and even some of them become professional photographers [1].

New users of digital single-lens reflex (DSLR) camera usually do not understand how to optimize the utilities of the camera. We can find that almost all DSLR cameras are equipped with auto-mode in taking a picture. The users do not need to set by their selves the exposure on the camera since this feature has already been set by a chip in the camera automatically. The presence of this auto-mode will certainly help the novice users to produce good photographs. Unfortunately, in a DSLR camera, manual-mode is precisely highlighted. Using features in the manual-mode, users can practice increasing their competency in photography. This manual-mode cannot be found in a pocket camera or in a camera of smartphone [1].

This research was intended to contribute in solving poor quality of pictures taken. Here we established an application that can evaluate taken photographs. The evaluation consists of suggestions
about how to improve the techniques of camera users in setting up speed, f-stop, ISO speed, and in applying flash and several tips and tricks to obtain good pictures. It is expected that users can better understand and improve their techniques in using their cameras from the suggestions given by the application.

Using the established application, it is expected that camera users can better understand techniques in utilizing their cameras. Since they can easily exercise and learn autodidact to improve their expertise in taking good pictures from the suggestions those given by the application. Hopefully, they can operate their DSLR cameras expertly, not only for prestige.

The rest of the paper is organized as follows. In Section 2, we discuss some basic concepts about brightness calculation estimation, face detection, exposure triangle, and forward chaining method which are used in this work. Section 3 consists of the research method performed. Section 4 provides the results of the research and the discussion about them. The conclusion is explained in Section 5.

2. Application overview

This application is established based on several methods in digital image processing [2]. The methods are used to sharpen pictures taken, to detect light intensity in pictures, and to detect the existence of faces in pictures. In this section, brightness calculation estimation, face detection, exposure triangle, and forward chaining method, which are used in our research, are explained.

2.1. Brightness calculation estimation

According to Wyszecki and Stiles [3], brightness is an attribute of visual sensation that are given in accordance with visual stimuli such as appearing brighter or less, or phenomena where the visual stimulus presented emits light with high or low intensity. Brightness is one of the factors that determine the quality of an image in determining the intensity of light in an image, just stepped or too dark. By using brightness calculation, a number is obtained which will then be analysed to determine whether the lighting is good enough or not.

The unit that is calculated for brightness of a photograph is luminance, L, as depicted by Equation 1, where:

\[ L = 0.27 r + 0.67 g + 0.06 b \]  

(1)

r (red), g (green), and b (blue) is the level of red, green, and blue, respectively, in a coordinate of RGB stimulus. Instead of the absolute luminance as in Equation 1, people usually use log average luminance \( L_{av} \) to judge the brightness of an image. The first mapping method involves scaling about some approximation of the neutral scene luminance or key of the scene. The log-average luminance is a good approximation of this and is defined as [5]:

\[ L_{av} = \exp \left( \frac{1}{\delta} \sum_{x,y} \log (\delta + L(x,y)) \right) \]  

(2)

The \( \delta \) value is a small bias included to allow log computations of pixels with zero luminance. The log average luminance is computed by summing the log-luminance of the pixel values of the image. This task can be approximated by sparsely sampling the image or operating on an appropriately resampled smaller version of the image. The latter can be accomplished using texture mapping operations to reduce the image size to \( 64 \times 64 \) before computing the sum of logs. If fragment programs are supported, the \( 64 \times 64 \) image can be converted to log-luminance directly, otherwise colour lookup tables can be used on the color values. The average of the \( 64 \times 64 \) log-luminance values is also computed using successive texture mapping operations to produce \( 16 \times 16, 4 \times 4, \) and \( 1 \times 1 \) images, finally computing the antilog of the \( 1 \times 1 \) image. Figure 1 shows some images with different values of log average luminance.
2.2. Face detection

Over the last few decades there have been several studies that have been developed and published relating to face detection. Here we explain several studies that have contributed widely to the development of face detection methods. Rowley et al. [6] have developed neural network based face detection. The technique used in this algorithm is to calculate the pyramid in the image to detect faces at different scales. Then, from the pyramid a sub window will be formed which will be corrected using non-uniform lighting and histogram equalization. The results will be further processed using several parallel neural networks that are useful for detecting faces. The results will be combined using logical AND, where logical AND serves to reduce the number of false face detection. In the early stages of development, this algorithm can only detect upright frontal faces. Another approach is by extracting database of many faces into several eigenfaces. An estimated face is reconstructed using linear combination of these eigenfaces. This method is called principal components analysis which also used in [7][8][9].

The next technique as used by Schneiderman and Kanade [10] is a statistical method for 3D object detection. The basic mechanism used by this algorithm is also by calculating the pyramid in the image. Then, scanning every sub window that has a fixed size in each layer of the pyramid. The contents of this sub window will be analysed using wavelet methods and histograms created for different wavelet coefficients. These coefficients will be tested on various trained parallel detectors that are sensitive to various types of orientation that represent an object. The orientation of this object is determined by the detector that produces the highest output.

Viola and Jones [11] carried out face detection using a multilevel classification. Each classification has features, the higher the classification, the more features will be. These features have been chosen to be able to know which parts of the image are part of the faces and not. Each classification eliminates parts of the image detected as no faces. Then, in the final classification, there will only be remaining portions of the image detected to contain faces. In this research, we use the method of Viola and Jones [11]. In the face detection, there are two types of errors that occur, namely, positive false (detecting faces in the image that actually do not have faces) or negative false (not detecting faces in the image that actually have faces).

The consideration of using the Viola-Jones method is because the Viola-Jones method has a high degree of accuracy compared to other methods. The researchers’ reports [11] on the percentage of error rates of the Viola-Jones method are as follows. The negative error rate (negative false) is under one percent and the positive error rate (positive false) is under 40 percent. These values are obtained when the framework only uses 32 filters to classify faces or not. Besides having a high degree of accuracy, the Viola and Jones method also does not take much time to do the process. This can occur because the Viola and Jones method does not analyse an image directly (per pixel) but instead analyses the image.
using features. The function of these features is to classify an image, an image that contains a face and an image that does not contain a face. There are three features based on the number of squares contained in it.

Calculating the values of these features is by subtracting pixel values in dark areas with pixel values in white areas. This method still seems to take a lot of time to calculate. Therefore, to speed up and simplify the calculation of the sum of the pixel intensities in these squares, a media in the form of an image integral is used. To get the value of this image integral, an accumulation of each square pixel is carried out from the top and left pixel value. After getting the value of the image integral at each point, it is now possible to calculate the value of the intensity in each square.

The calculation of the value of this feature is done sequentially from the leftmost side of the image to the bottom right with a shift of $\Delta x$ and $\Delta y$. The smaller the value of $\Delta x$ and $\Delta y$, the detection results will be more accurate. But the trade-off is a longer process time.

After doing these calculations, the features and their square shapes are tested using a set of images containing faces and a set of images that do not contain faces with an image size of 24 x 24. Then, a routine finds out which features can recognize faces and discard images that do not have faces. The Viola and Jones method performs this routine by using a database from the internet consisting of 4000 images containing faces and 10,000 images without faces. By using these 24 x 24 images, there are 45,000 different ways to implement three types of features in the images. With this number of possibilities which is very far compared to the number of pixels present in a 24 x 24 image (576 pixels). With this number of possibilities, it is impossible to calculate for all features, therefore it is necessary to reduce the number of possibilities. Ada-Boost algorithm is used to choose which features will be removed, so only certain features will be included.

The function of the Ada-Boost algorithm is to search for features that have a high level of differentiation. This is done by evaluating each feature based on training data. Through values of the features, it can be selected features which have the biggest threshold between faces and non faces, which will be considered as the best features. From this classification done by Ada-Boost, two features are found that have a positive error rate of 40 percent, in other words these features will remove 60 percent of images that do not contain faces. An instance of classification in 24x24 image which is performed by each feature, is shown in Figure 2.

![Figure 2. Classification Example in a 24 x 24 Image using Two Best Features][9]

Figure 2b shows the horizontal feature whose task is to measure the difference in light intensity in the darker areas of the eye compared to the brighter areas of the cheekbones. Then, the feature in Figure 2c takes measurement of the two darker areas of the eye when compared to the nasal bone area. Viola-Jones algorithm uses multilevel classification. For this study, a classification which has 3 levels is used, where each level is tasked to exclude sub-images that are detected non faces. This is done because it will be easier to analyse sub images that do not contain faces than to analyse sub images that contain faces. These steps can be visualised into Figure 3. In fact the multilevel classification developed by Viola and Jones uses 32 classification levels and 4297 features.
2.3. Exposure triangle
In the world of photography, the term exposure triangle is certainly important. Exposure triangle is three factors in photography that determine how much the intensity of light in an image. The concept of the exposure triangle was first popularized by Bryan Peterson through his book entitled "Understanding Exposure" [12]. In this book, three factors are explained that are ISO speed, Shutter Speed, and Aperture [13][14]. Each factor has the same weight in influencing exposure on the image. Figure 4 shows a clearer description of the role of each factor.

2.4. Forward Chaining
The method used here to provide advice to users is the expert system of forward chaining. This method works by tracing ahead from a set of facts to look for hypotheses that can be used as conclusion [15]. This forward chaining is sometimes also called data-driven method, this is because the inference engine used is information that is already available. From this information, conclusion will be obtained from a problem. In this research, the available information is exif data from each image and input from the user in the form of information about the location of the image taken (outdoor/indoor).

3. Research method
In this research, application performance testing is conducted to examine an image until it can provide an accurate decision or suggestion in accordance with the actual image condition. This research uses a quantitative approach, by calculating, measuring and comparing data of actual image conditions. This calculation process is useful for obtaining blur values and luminance values that can affect the quality of the pictures. We want to find out how effective this application can help people in learning photography.

The first step is to determine which factors can affect the quality of photographs taken using a digital camera. Subsequently an assessment of the quality of the photograph is done using these factors. To provide the good recommendations, the system checks the specifications used when taking pictures namely shutter speed, f-stop, iso speed, and focal length.

Overall, this application uses 3 main functions, namely blur detection, brightness calculation estimation, and face detection. These functions serve to provide an assessment of the factors that generally affect the quality of the images. After the system knows the image quality, the next step is to
check the exchangeable image file format data (exif data) that is in the image. Exif data is similar to metadata whose job is to store important information about a file. However, this exif data is more devoted to digital image files.

Then, from a lot of information in exif data, we only use some information to check, i.e. shutter speed, ISO speed, f-stop, and time created. With this information, the application can analyse and find out what causes the images to be of poor quality. After that, the application can provide suggestions to users in accordance with the deficiencies contained in the way the image is taken. Figure 5 shows steps that need to be taken to use the application correctly. The system’s workflow can be seen in the flow chart of Figure 6.

Figure 5. General description of the system

Figure 6. System’s flowchart
The first thing a user must do is to press the browse button to select the desired image. Then, press the OK button. The image that the user chooses will appear in the picture box column. To start the analysis process, press the analyse button. When the analyse button is pressed, the first instruction that is executed is to retrieve the exif data in the image, this exif data is extracted using the levdan library. The next step is to look for the log average luminance, $L_{avg}$, value, this $L_{avg}$ value will be an indicator of light intensity in the image, this $L_{avg}$ will be stored in a variable. Then, proceed to the blur detection process, the face detection process will be performed. When the application detects the presence of a face the blur value will be averaged using that face area. If the application does not detect a face, the blur value used is based on the entire image area. After all the required data has been obtained, the application will determine the image quality based on the parameters that have been obtained. From this image quality, the application will give suggestions to improve a poor image and the application also tells if the quality of the picture taken is good enough.

4. Results and discussion

4.1. Results of stage one experiment

At stage one, the experiment is performed using sample-1 which consists of 47 images. From 47 tested images, there are 10 images that the analysis of their conditions is not match with their actual conditions. So according to the test results in sample-1, the level of accuracy that can be given by the application is 78.7%. This 78.7% value is obtained from a combination of errors that occur in blur detection and luminance calculation. There are 8 blur values and 3 $L_{avg}$ values that are not match with their actual conditions. Blur detection itself has an accuracy rate of 82.98%. While the accuracy rate on the luminance calculation function is quite high at 93.62%. Testing is also done for sample-2 where the number of images tested in sample-2 is 45. The analysis error is found in 11 images. So based on the results of testing in sample-2, the level of accuracy provided by the application is 75.5%. Just like the previous experiment, this accuracy value is obtained from the total errors that occur in the blur detection and luminance calculation. Blur detection has an accuracy rate of 88.89%, the luminance calculation itself has an accuracy rate of 84.44%. From testing on the two types of samples above, it can be concluded that the accuracy of the analysis in the application is in the range of 75% - 79% and most of the errors that occur due to the value of the lumen / blur in the image that is slightly above or below the threshold value.

4.2. Results of stage two experiment

In the stage 2 experiment, several images will be selected from the sample-1 and tested whether suggestions given can improve the images. In first test, a sharp image in Figure 7 is used, but with an over exposure condition. The shooting location is outdoors. The suggestion given by the application to improve the image is to increase the value of the f-stop several steps. Result of first test can be seen in Figure 8.

![Figure 7. Application display in first test](image-url)
The resulted image after following the suggestions of the application in first test to increase the f-stop value can be seen in Figure 8. It can be seen in that the image is getting better, no longer over-exposure like the previous one after the f-stop is increased from 2.8 to 5.6. In the second test, an image with low light conditions were used. The application provides information that shooting during the day rarely lacks light except for cloudy weather. Then the application suggests reducing the shutter speed or f-stop value. Figure 9 shows the resulted image after following the suggestion from the application, i.e. by changing shutter speed from 1/1000 to 1/125, the image doesn't look dark or under exposure anymore.

In third test, without using the face detection feature, the application provides an assessment that the image is blurry and over exposure. Then, the face detection feature is turned on, the result of the application analysis turns into only an over exposure error. The application provides suggestion for increasing the shutter speed value or by reducing the ISO value. After the suggestion is followed, the image looks much better than before, with the right light intensity and sharpness is also good by increasing the shutter speed from 1/20 to 1/100, as seen in Figure 10. In fourth test, the application

**Figure 8.** Resulted image from first test after increasing f-stop from 2.8 to 5.6

**Figure 9.** Resulted image from second test after decreasing shutter speed from 1/1000 to 1/125

**Figure 10.** Resulted image from third test after increasing shutter speed from 1/1000 to 1/125
evaluates the image under the correct exposure condition but too blurry. Even though the actual condition is that the image is sharp but over exposure. Because this analysis error causes suggestion given by the application to be less precise. The suggestion given is to increase the f-stop value, but the application suggests increasing the shutter speed. In fifth test, in taking an image with a moving object, the application will suggest raising the shutter speed value in the range of 1 / 125s - 1 / 500s. With a shutter speed high enough, the moving object can be easily frozen.

In sixth test, we used an image with blur and over exposure condition. Blur condition in this image can be caused by image that are too bright resulting in details in the image are not visible. The suggestion given by the application is to increase the shutter speed value and ensure the focus point is in the right position. In seventh test, an image was used with the condition of over exposure and blur taken indoors. The application estimates that this deficiency is caused by setting the shutter speed value too low or using the wrong external flash.

4.3. Validation of the application’s suggestions

To find out how well the application's suggestions are given to users, validation is done to professional photographers. It is expected to know whether the suggestions given by the application is in accordance with the opinions of a professional photographer. Validation is done by Mr. Denny Hendrata from Voi & Vox, where the validation was given in the form of his comments about the images by looking at the conditions in the images and the exif data listed without given suggestions from the application. This kind of validation is known as a blind validation. In this way the professional photographer can provide comments in accordance with his/her knowledge that are not affected by the suggestions given by the application.

Validation is done with a total of 10 random images. Out of the 10 images, there are 7 images that have matching suggestions of the application and comments of an expert. This is partly because there is an error in the detection of images’ qualities by the application. This causes the suggestions of the application do not match the comments given by photography expert. He has a higher tolerance for a condition, such as over light condition which sometimes can still be tolerated. From the validation results, it can be concluded that the application is good enough in providing suggestions.

5. Conclusion

From the results of experiments on the application there are several conditions in which the application made a mistake in assessing the image. From several samples conducted in the experiments, it is obtained an average level of accuracy of 75%. This error occurs in a combination of one or two conditions, this condition is when blur detection and analysing the light intensity condition in the image. The blind validation done by an expert has resulted in 70% matching image analyses by the applications compared to comments of the expert.

6. References

[1] Ang T and Burgeon V 2016 Fundamentals of Photography (London, UK: Octopus Publishing Group Ltd.) 104-131
[2] Gonzales R C and Woods R E 2018 Digital Image Processing 4th Edition (Essex, England: Pearson Education Ltd.) 48-54
[3] Wyszecki G and Stiles W S 2000 Color Science: Concepts and Methods, Quantitative Data and Formulæ 2nd Edition (New York, USA: John Wiley & Sons)
[4] Chaves J 2016 Introduction to Nonimaging Optics 2nd Edition (Florida, USA: CRC Press) 679
[5] McReynolds T and Blythe D 2005 Image Processing Techniques: Advanced Graphics Programming Using OpenGL (San Francisco, USA: The Morgan Kaufmann Publishers) 213
[6] Rowley H A, Baluja S and Kanade T 1996 Neural network-based face detection IEEE Trans. on Pattern Analysis and Machine Intelligence 20(1) 203-208
[7] Hugeng H, Wahab W and Gunawan D 2015 A new selection method of anthropometric parameters in individualizing head-related impulse responses TELKOMNIKA 13(3) 1014-1020

[8] Hugeng H, Anggara J and Gunawan D 2017 Implementation of 3D HRTF interpolation in synthesizing virtual 3D moving sound Int. J. of Technology (IJTech) 8(1) 186-195

[9] Hugeng H, Anggara J and Gunawan D 2017 Enhanced three-dimensional HRIRs interpolation for virtual auditory space Proc. IEEE Int. Conf. on Signals and Systems 2017 (Bali, Indonesia, May 16-18) 35-39

[10] Schneiderman H and Kanade T 2000 A statistical method for 3D object detection applied to faces and cars Proc. IEEE Conf. on Computer Vision and Pattern Recognition (South Carolina, USA, June 2000) 746-751

[11] Viola P and Jones M J 2000 Robust real-time face detection Int. J. of Computer Vision 57(2) 137-154

[12] Peterson B F 2004 Understanding Exposure: How to Shoot Great Photographs with Film or Digital Camera (New York, USA: Amphoto Books)

[13] Sheppard R 2010 Digital Photography: Top 100 Simplified Tips & Tricks 4th Edition (Indianapolis, USA: John Wiley and Sons) 40

[14] Rashidian Vaziri M R 2015 Role of the aperture in Z-scan experiments: A parametric study Chinese Physics B. 24(11) 114206

[15] Al-Ajlan A 2015 The comparison between forward and backward chaining Int. J. Machine Learning and Computing 5(2) 106-113