Measuring Similarity Analysis for Image Verification

Ho Ryu[1*] and Il-Young Moon[2]

[1,2]Department of CSE, KOREATECH University, Cheon-an, Republic of Korea
[1]r7826974@gmail.com, [2]iymoon@koreatech.ac.kr

Abstract

In this paper, we try to find a new decision method to solve the problem of blurred objects in the image for probabilistic approach. Sometimes we have difficulties to restore the details in the low sampling quality image. At first, we will classify into two groups as accurate and variable reference environment. And we will also consider one more option. It's the digital image with no reference condition. It will be useful to find special features in the insufficient information in the poor digital image environment.

1. Introduction

Recently digital image verification is main area of the digital forensics. In the digital forensics, we should extract hidden information in the digital image. Digital image is mapped to the pixel from the natural image by digital device [1]. If the resolution is poor, the digital image should be processed to gain the information. The most important thing in the digital forensic area is to find information from the digital image. In this paper, we will classify into the three groups by the next reference environment conditions [2].

2. System Model and Methods

In Figure 1, we use some digital image processing methods for probabilistic approach. We can classify digital image by the next three ways.

Figure 1. Digital Image Processing Method for Probabilistic Approach

2.1. Measuring Similarity with Accurate Reference Environment

The accurate reference image indicates the objective certain image. The car registration plate is a good sample to find out certain number. It means that we can get accurate image by calculating the probability about a certain number. The logical function of digital image pixel can measure the probability with the binary image and reference image.

$$\text{Similarity} = \text{OR logical extracted image pixel / Reference image pixel} \qquad (1)$$

In the equation (1), the reference image indicates the objective number from the car registration plate. It means we can get accurate image area to calculate the probability of a certain number. The OR logical function means the numerator from the binary image and reference image in equation (1). Figure 2 shows the principle about similarity in the accurate reference environment.
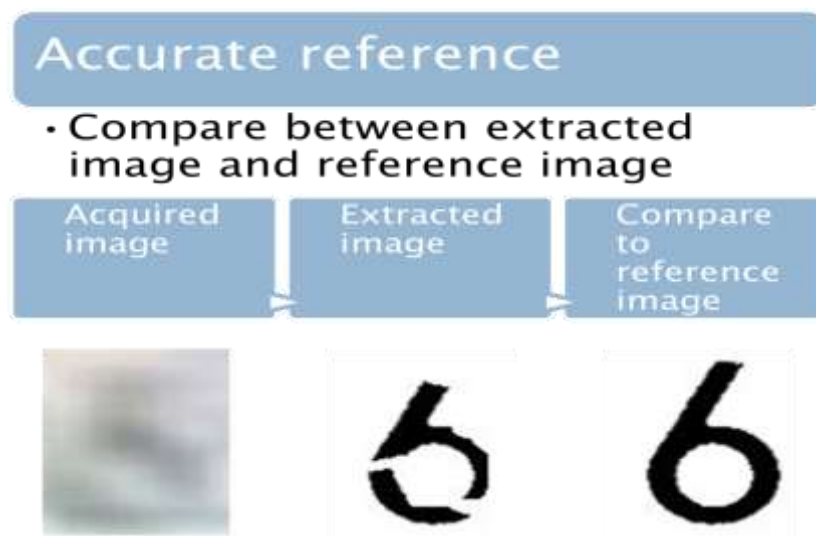

Figure 2. Measuring Similarity with Accurate Reference Environment

2.2. Measuring Similarity with Variable Reference Environment

The variable reference image is used for the case that the reference is not exact. We calculate the similarity of two images as following equation.

$$\text{Similarity} = \text{OR logical extracted image pixel / Montage image pixel} \qquad (2)$$

In the equation (2), the montage image is consisted with vague acquired image and some possible reference [3]. It means we don't know the exact information about the reference and just assume about curious reference whether it is right or not [4]. Then we can get the similarity between montage and acquired image. The OR logical function means the numerator from the binary image and montage image in equation (2). Figure 3 shows the principle about similarity in the variable reference environment. In the color image, to get the objects, we should consider the visual attention not only the luminance but also the color [7-8].
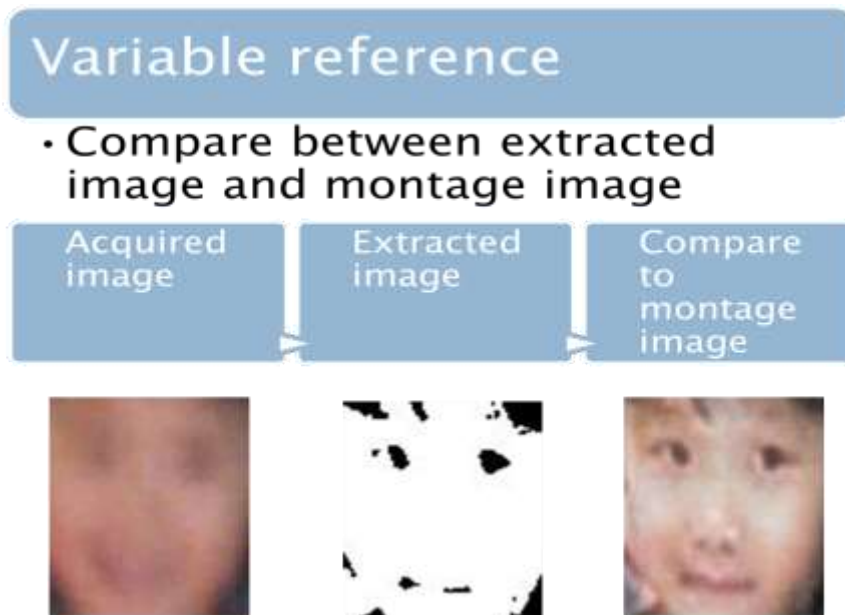
Figure 3. Measuring Similarity with Variable Reference Environment

2.3. Measuring Similarity with No Reference Environment

In the case of no reference image indicates that we don't have any reference image to compare. If we don't have reference image, it is impossible to decide that the certain work of art is genuine or not. Especially plagiarism debates are still exist in contemporary arts in Korea [5]. The essential reason of controversial in art works due to insufficient evidence to compare. Technically, the only way to find the genuineness is examining other paintings of certain artist. In this paper, we analyze the habit of painter for the analysis certain paintings by statistics way about the art images. Figure 4 shows correlation of the paintings about the peculiarity of one painter.
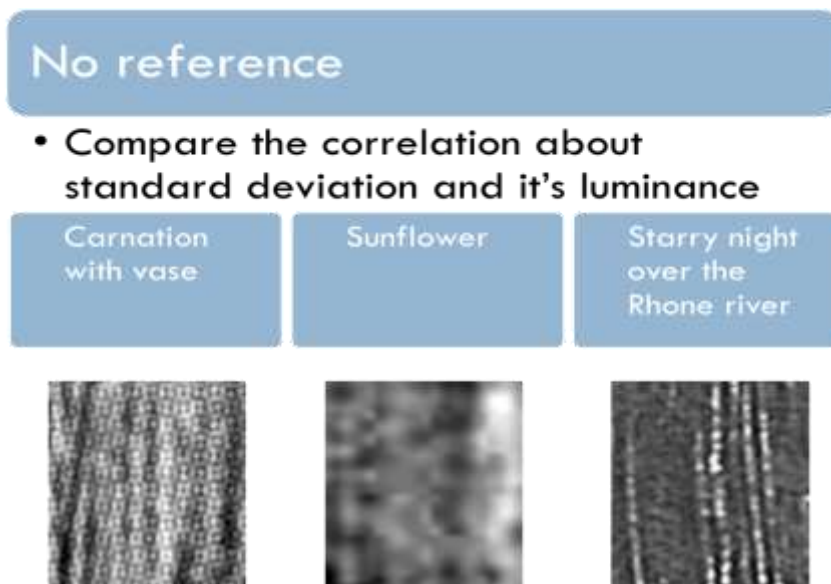

Figure 4. Measuring Characteristic Points of One Painter Art

3. Results

3.1. Measuring Similarity with Accurate Reference Environment

We adapt for car registration number to prove validation of theory. Then we measure similarity analysis by the probabilistic method in this thesis. Figure 5 shows the process to get the extracted image by accurate reference image.

Figure 5. Process to get the Extracted Image by Accurate Reference Image

3.2. Measuring Similarity with Variable Reference Environment

The montage image is consisted with vague acquired image and some possible reference. It means we don't know the exact information about the reference and just assume about curious reference whether it is right or not. Then we can get the similarity between montage and acquired image. Figure 6 shows the process to get the extracted image by variable reference image.

Figure 6. Process to get the Extracted Image by Variable Reference Image

3.3. Measuring Similarity with No Reference Environment

We also consider one more option. It's the digital image with no reference condition. We analyze standard deviation with luminance in the peculiar point in the picture. The mean value of luminance and standard deviation value of special feature will be a clue to investigate the art image. Figure 7 shows correlation between standard deviation and luminance in the image.
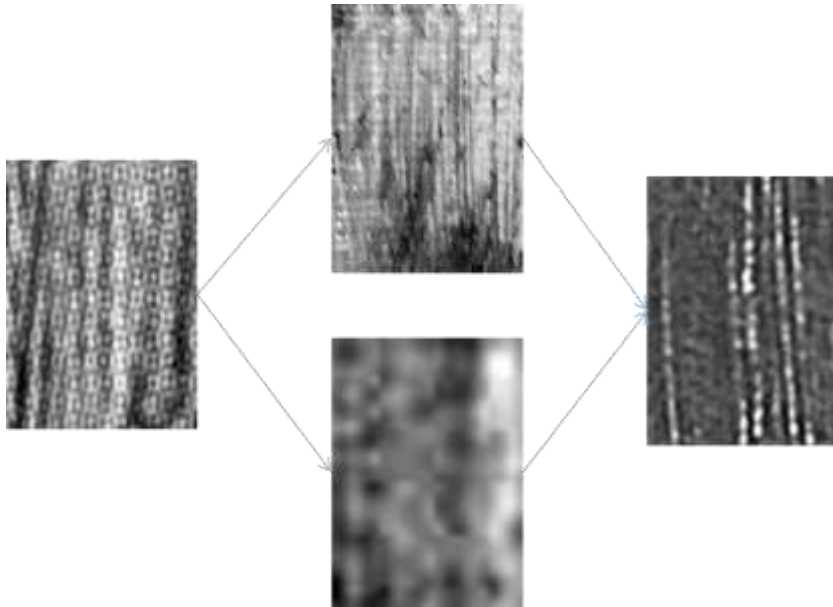

Figure 7. Correlation between Standard Deviation and Luminance

Figure 8 shows the histogram of acquired painting image about 'carnation with vase' by Van Gogh. Figure 9 shows the histogram after processing about 'carnation with vase' for the measuring similarity each painting.
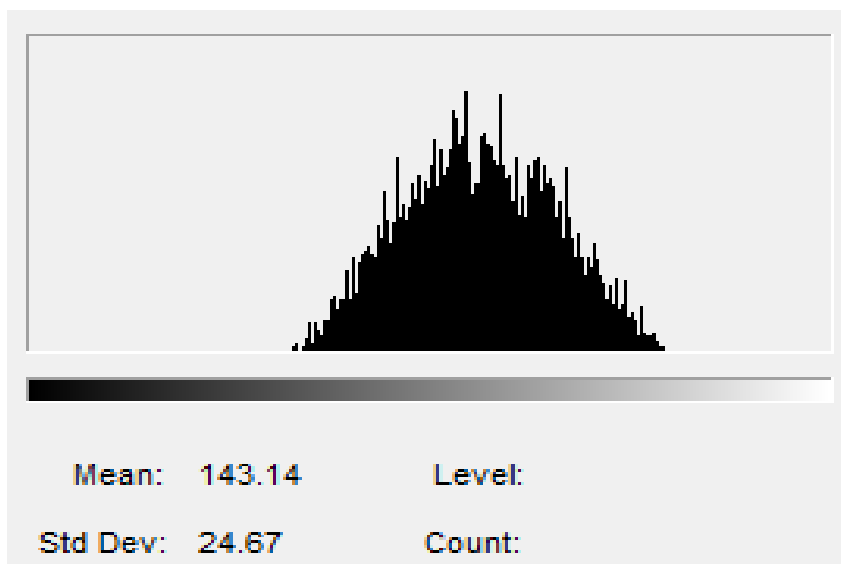

Mean: 143.14   Level:

Std Dev: 24.67   Count:

Figure 8. The Histogram of Acquired Image about 'carnation with vase'

In Figure 8, the acquired image has essential problem to find out the information. It shows the blurred image due to the poor environment to acquire. Normal distributed

image shows the concept of probability density function for the experiments in Figure 9. The histogram stretching changes the level of contrast in the digital image as a normal distribution. It shows the concept of probability density function.
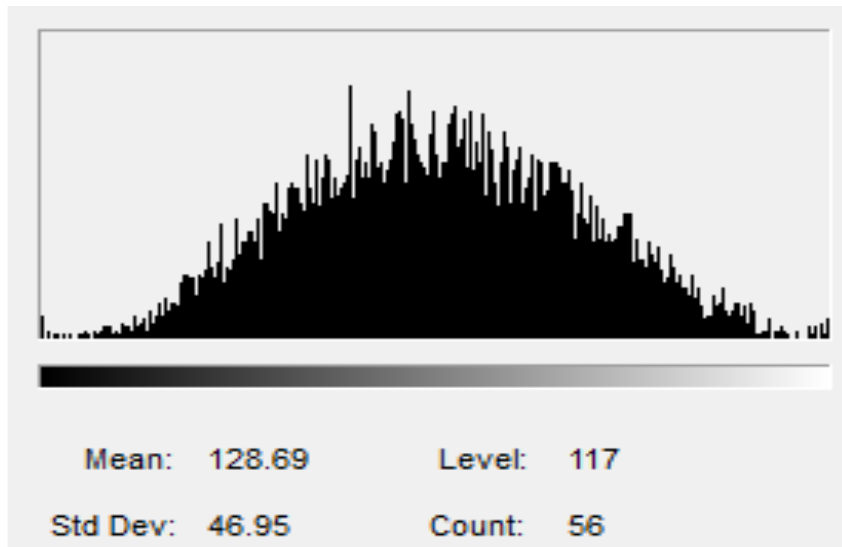

Mean: 128.69   Level: 117

Std Dev: 46.95   Count: 56

Figure 9. After Processed Image about 'Carnation with vase'

The output image will be changed as following function. The degree of intensity about each pixel, $P'(x, y)$ is intensity value after the normal distribution processing.

$$P'(x, y) = \frac{p(x, y) - \min}{\max - \min} \times 256,$$ (3.1)

Where P(x, y) is the intensity value of every each pixel about the acquired image.

The output image will be changed as following function. The degree of intensity about each pixel, $P'(x, y)$ is intensity value after the normal distribution processing.
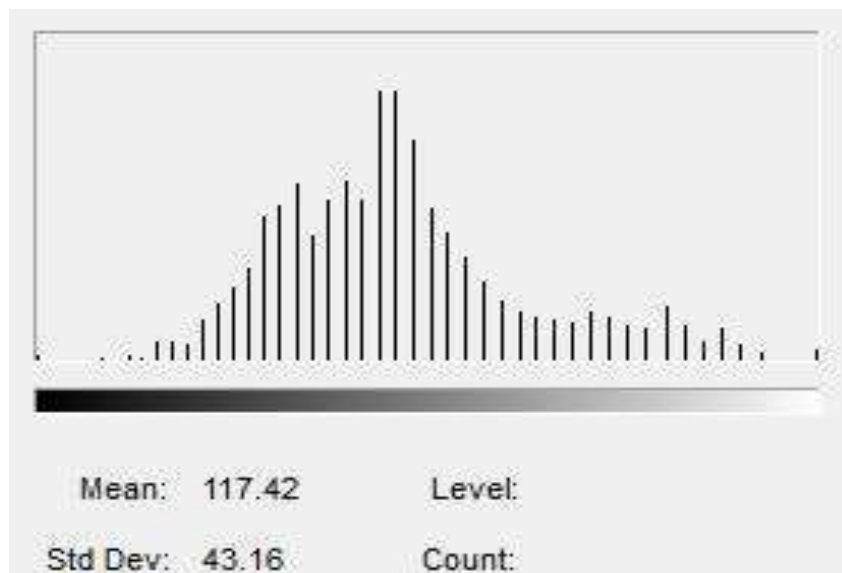

Mean: 117.42   Level:

Std Dev: 43.16   Count:

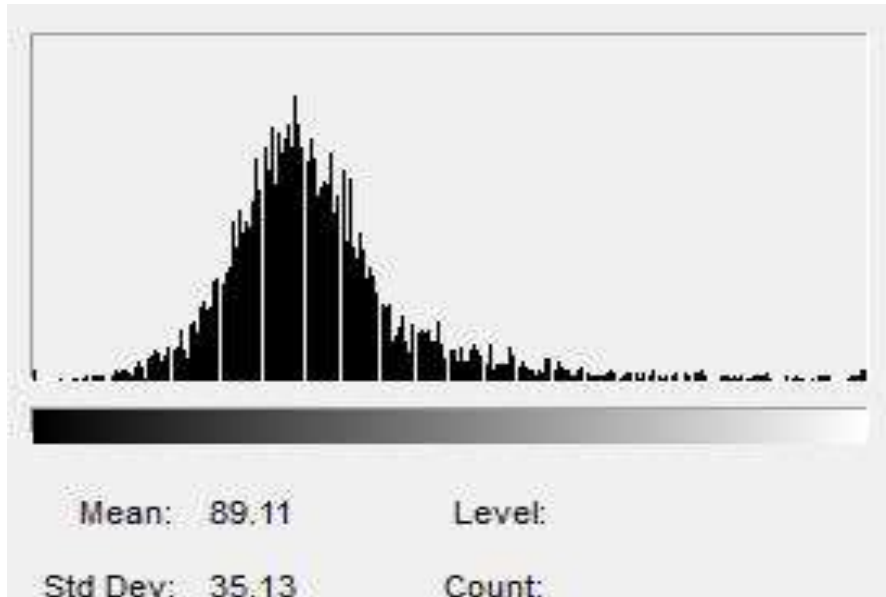**Figure 10. After Processed Image about 'Sunflower'**

Figure 11. After Processed Image about 'Rhone river'

To find out regression character, the other peculiarity points will be measured in the same painter for making one order polynomial. Figure 10 and Figure 11 show the characteristics for search regression discontinuity design.

4. Conclusions

4.1. Similarity Analysis with Accurate Reference Image

Table 1. Measuring Similarity in Car Registration Numbers

Reference Image	Extraction Image	Similarity	Reference Image	Extraction Image	Similarity
0		0.656	5		0.589
1		0.665	6		**0.751**
2		0.490	7		0.490
3		0.645	8		0.575
4		**0.725**	9		0.465

In the experiments in this paper, we adapt for car registration number to prove validation of theory for accurate reference. If the reference is exact, we calculate similarity with image based on the reference. We measure similarity analysis by the

probabilistic method in this thesis. The results show the similarity between each reference and extraction image. In the Table 1 and Figure 12, number 6 and 4 show the strong similarity between reference numbers.

Figure 12. Binary Image after Processing and Strong Association Numbers

4.2. Similarity Analysis with Variable Reference Image

In the experiments in this paper, we adapt for vague face image to prove validation of theory for variable reference. If the reference is not exact, we calculate similarity of two images after making the montage image.

In the Table 2, the montage image is consisted with vague acquired image and some possible reference. It means we don't know the exact information about the reference and just assume about curious reference whether it is right or not. Then we can get the similarity between montage and acquired image.

Table 2. Measuring Similarity for Vague Face

Montage Image	Extraction Image	Similarity
		0.667
		0.751

4.3. Similarity Analysis with No Reference Image

We also consider one more option. It's the digital image with no reference condition. In this paper, mean value and standard deviation of special feature is used for investigate the art image.

We used standard deviation and difference between Y value and Mean value(Y=127.5). With the two statistical value, we can draw one order polynomial by regression discontinuity design [6]. Figure 13 shows characters about the two painting arts comparison with statistical method.
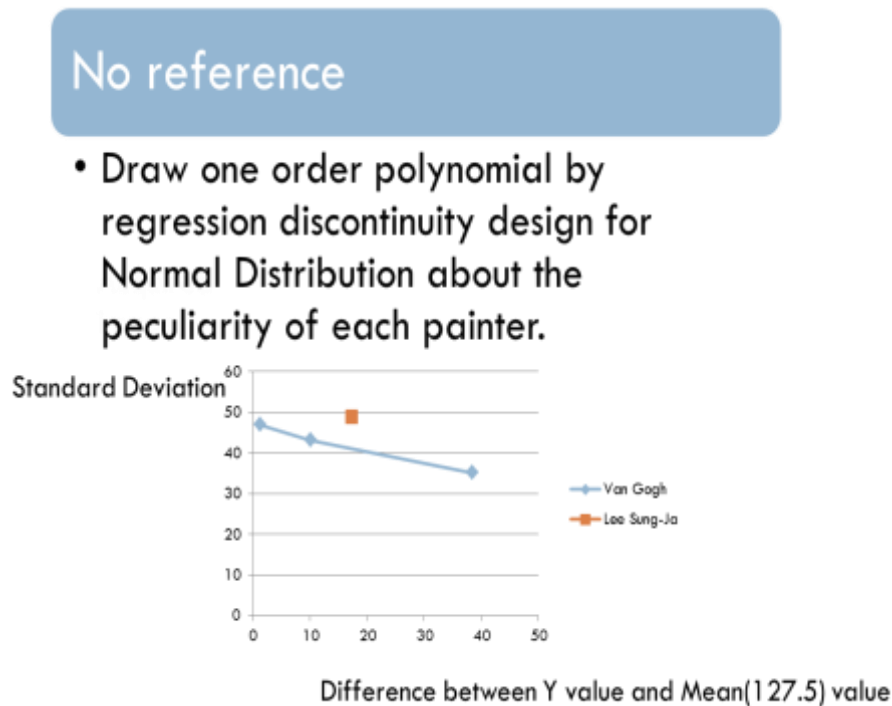

Figure 13. Character Comparison with Statistical Method

Acknowledgments

References

[1] D. L. Ruderman and W. Bialek, "Statistics of Natural Images: Scaling in the Woods' Physical Review Letters", vol. 73, no. 6, (**1994**) August 8.
[2] A. Piva, "An overview on image forensics", ISRN Signal Processing, (**2013**).
[3] R. Uthayakumar and A. Gowrisankar, "Generalized Fractal Dimensions in Image Thresholding".
[4] S. S. Bedi and R. Khandelwal, "Various image enhancement techniques-a critical review", International Journal of Advanced Research in Computer and Communication Engineering, vol. 2, no. 3, (**2013**).
[5] E.-s. Oh, "A Study on the Plagiarism Debates in Contemporary Art: The Case of Mongjoo Son vs. Jeonghyun Park", The Journal of Art Theory & Practice, vol. 23, (**2017**), pp. 152-188.
[6] A. Gelman and G. Imbens, "Why high-order polynomials should not be used in regression discontinuity designs", (**2017**).
[7] S.-G. Cho, J.-J. Hwang and N.-J. Kwak, "Visible distortion predictors based on visual attention in color images", Journal of information and communication convergence engineering, vol. 10, no. 3, (**2012**), pp. 300-306.
[8] Y.-S. Moon, S.-P. Choi, E.-K. Lee, J.-J. Kim and H.-R. Choi, "Performance Evaluation of Advanced Container Security Device(ACSD) system based on IoT(Internet of Things)", Journal of the Korea Institute of Information and Communication Engineering, vol. 17, no. 9, (**2013**), pp. 2183-2190.

Authors

Ho Ryu, Department of Computer Science Engineering,
Korea University of Technology & Education
E-mail: r7826974@gmail.com

Il-young Moon, Department of Computer Science Engineering,
Korea University of Technology & Education,
E-mail: iymoon@koreatech.ac.kr