

Regression Analysis Based Variance Estimation of Gaussian Distribution for Histogram Matching

Yusuke Kawakami ^{1*}, Tetsuo Hattori ², Yoshiro Imai ², Kazuaki Ando ², Yo Horikawa ², R. P. C. Janaka Rajapakse ³

¹ DynaxT Co., Ltd.,
2217-6 Hayashi

Takamatsu City, Kagawa 761-0301, Japan
E-mail: riverjp2002@gmail.com

² Graduate School of Kagawa University,
2217-20 Hayashi

Takamatsu City, Kagawa 761-0396, Japan
E-mail: hattori@pe.kagawa-u.ac.jp, {imai, ando, horikawa}@eng.kagawa-u.ac.jp

³ Tainan National University of the Arts,
66 Daci

Guantian District, Tainan 72045, Taiwan
E-mail: janakaraja@gmail.com

Abstract

This paper describes an improved method of the variance estimation of Gaussian distribution for Histogram Matching based on Gaussian Distribution (HMGD). In our previous paper, focusing on the symmetry of the Gaussian function, we presented another method for the variance (or Gaussian width) estimation of Gaussian distribution in the original histogram. However, since the real shape of mountain like Gaussian function in the original image's histogram does not always show the good symmetry, the variance estimation method that we previously presented did not work so well as we expected. In this paper, we newly propose the improved estimation method using regression analysis, based on curvature computation for the cumulative histogram of original image's one. In the newly proposed method, first, we detect the histogram peak of original image's histogram by using curvature computation; next, we perform the regression analysis for the cumulative histogram, using an approximated function of the curvature that includes the variance parameter. Also in this paper, we show some experimental results by the estimation method.

Keywords: Image processing, Cumulative histogram, Curvature, Variance estimation, Histogram matching

1. Introduction

Recently, automated image processing for enhancement of color images has been familiar in various equipment, for example, Digital Signage, Smart Phone, etc. ¹⁻³

In our previous paper, we presented a Histogram Matching based on Gaussian Distribution (HMGD) that is one of the automated image arrangements processing method using Elastic Transformation⁴⁻⁵ on the brightness axis. And we have illustrated that the HMGD

* 1882-36, Yashima-nishi, Takamatsu City, Kagawa 761-0113 Japan

processing could improve the feeling (or Kansei) impression better than the original image.⁶

In order to further develop the HMGD processing method for better automated one, we noticed that the variance estimation of Gaussian distribution around the neighborhood of peak in the original image's histogram should be performed. Then previously, we presented a variance (or Gaussian width) estimation method using the symmetry of Gaussian function.

However, since the peak shape like Gaussian function in the histogram of original image does not always show the good symmetry in real, the previous variance estimation method did not work so well as we expected. In this paper, we newly propose the improved estimation method using regression analysis, based on curvature computation for the cumulative histogram of original image's one. Moreover, we show the experimental results.

2. Principle

2.1. Brightness Peak Detection of Original Image

In the section, we describe the principle of peak detection in the brightness (or lightness) histogram of original color image.

Let y be a function with respect to x that represents the cumulative histogram curve and the brightness parameter, respectively. The definition of the curvature $R(x)$ for the cumulative histogram curve is given by the following Eq. (1).⁶⁻⁹

$$R(x) = \frac{d^2 y}{dx^2} \bigg/ \left(1 + \left(\frac{dy}{dx} \right)^2 \right)^{\frac{3}{2}}. \quad (1)$$

And let $g(x)$ and K be a Gaussian distribution function and a coefficient which is defined by the following equation, respectively.

$$g(x) = \frac{K}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right), \quad (2)$$

$$\frac{K}{\sqrt{2\pi\sigma^2}} \int_0^L \exp\left(-\frac{(u-a)^2}{2\sigma^2}\right) du = 1.$$

Next, let $y=f(x)$ be a function that representing the cumulative histogram curve that is represented Eq. (3). That is, dy/dx and d^2y/dx^2 be described as Eq.(4) and

Eq.(5), respectively. From Eq.(4) and Eq.(5), we obtain the curvature $R(x)$ as shown in Eq. (6).

$$f(x) = \int_0^x g(u) du = \frac{K}{\sqrt{2\pi\sigma^2}} \int_0^x \exp\left(-\frac{(u-a)^2}{2\sigma^2}\right) du. \quad (3)$$

$$\frac{dy}{dx} = g(x) = \frac{K}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right) \quad (4)$$

$$\frac{d^2 y}{dx^2} = \frac{dg(x)}{dx} = \frac{(a-x)}{\sigma^2} g(x) \quad (5)$$

$$R(x) = \frac{\frac{(a-x)}{\sigma^2} g(x)}{\left(1 + g(x)^2\right)^{\frac{3}{2}}} = \frac{(a-x)}{\sigma^2} \left(\frac{g(x)}{\left(1 + g(x)^2\right)^{\frac{3}{2}}} \right) \quad (6)$$

From Eq.(6), since the value of $g(x)$ is greater than or equal to 0, we understand that the curvature $R(x)$ varies the sign according to the variable x .⁹

That is, if $x < a \rightarrow R > 0$ (downward convex shape), and if $x > a \rightarrow R < 0$ (upward convex shape).

2.2. Variance Estimation

The conceptual image of the original image's histogram with the variance σ^2 and the average a is shown in Fig.1.

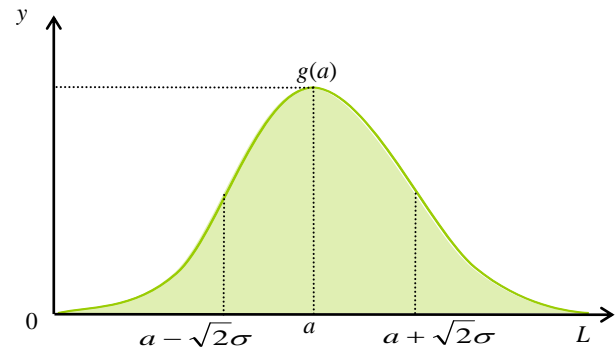


Fig.1. Conceptual histogram image of the Gaussian function curve around the peak in the histogram.^{8,9}

From Eq. (6), we can also represent $R(x)$ as Eq. (7).

$$R(x) = \frac{\frac{(a-x)}{\sigma^2} g(x)}{\left(1 + g(x)^2\right)^{\frac{3}{2}}} = \frac{1}{\sigma^2} \left(\frac{(a-x) g(x)}{\left(1 + g(x)^2\right)^{\frac{3}{2}}} \right) \quad (7)$$

Let $C=1/\sigma^2$ and let the other function excepting the C in Eq. (7) be $H(x)$, then we have the following Eq. (8).

$$R(x) = \frac{1}{\sigma^2} \left(\frac{(a-x)g(x)}{(1+g(x)^2)^{\frac{3}{2}}} \right) = CH(x) \quad (8)$$

From Eq. (8), we obtain relation (Eq. (9)) on sampled data where ε_i shows the error.

$$R(x_i) = CH(x_i) + \varepsilon_i, \quad \varepsilon_i \sim N(0, \rho^2) \quad (i=1, \dots, n) \quad (9)$$

Applying the least square regression analysis method to the sample data, we have Eq. (10).¹⁰

$$C = \frac{\sum_{i=1}^n (H(x_i) \cdot R(x_i))}{\sum_{i=1}^n (H(x_i))^2}, \quad \sigma^2 = \frac{1}{C} \quad (10)$$

However, for the computational simplicity, we roughly approximate the $H(x)$ as shown in Eq. (11).

$$H(x) = \frac{(a-x)g(x)}{(1+g(x)^2)^{\frac{3}{2}}} \approx (a-x)g(x) \quad (11)$$

3. Experiment Results

We have done the experimentation in order to evaluate the capability of the variance (σ^2) estimation over the sampled data where the number of the sampling points is s . By varying the number s , we performed the estimation and then applied the HMGD processing to the original image (Fig.2) based on the estimated variance. We initiated the number of sampling points s as $s=5$ and varied the number increasing by 4, such as $s=5, 9, 13, 17, 21, 25$.

Experimental results are shown in Fig.3 through Fig.8. In each figure, the color image is located at left upper side and the corresponding gray level image (= image showing intensity) at left lower side. Moreover, the histogram of the gray level image is shown at the right upper side, and its cumulative histogram at the right lower side.

Also in each figure, HMGD processed images (color and gray level image), the histogram corresponding to

the gray level image, and its cumulative histogram are illustrated with the number s and the estimated standard deviation (STD) σ .

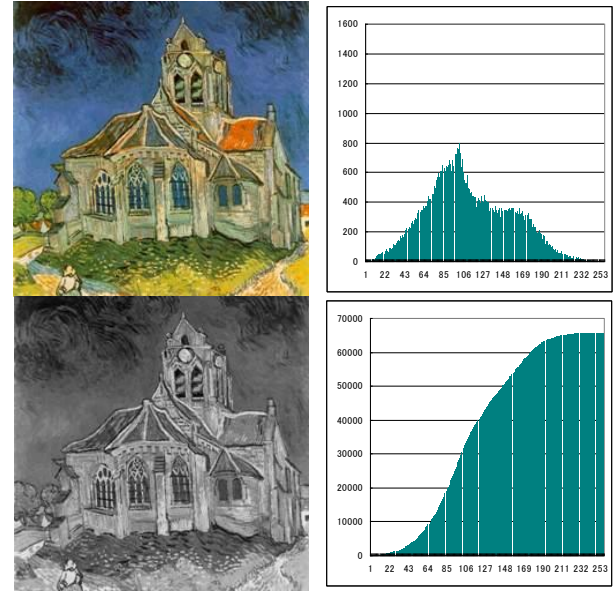
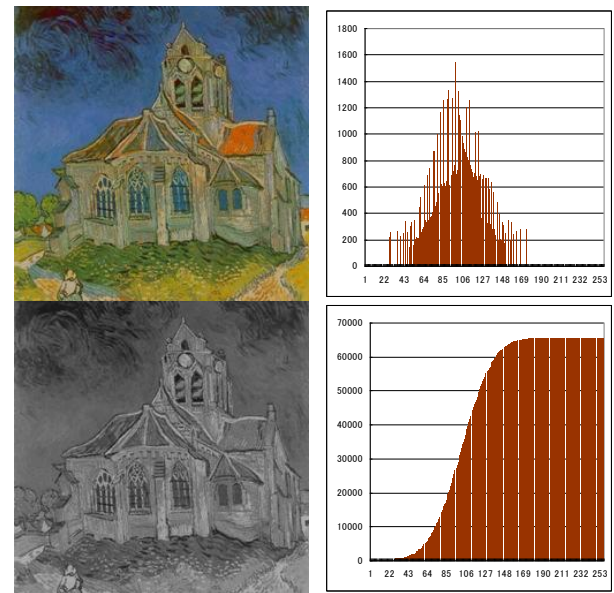
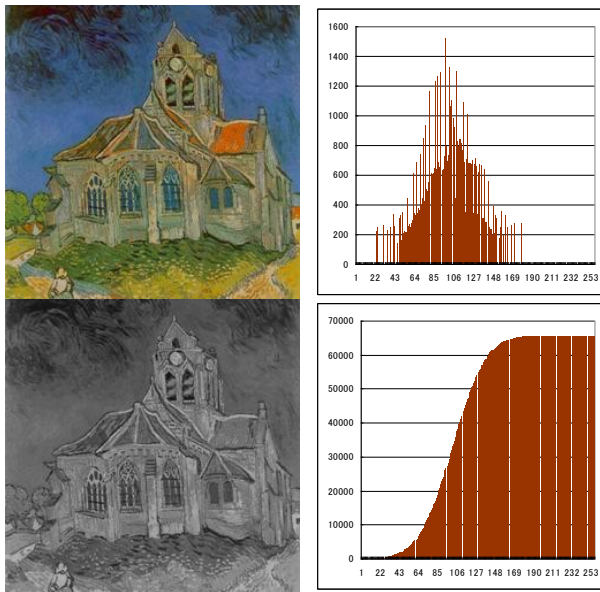


Fig. 2. Original images (color and gray level) and the histogram of the gray level image and its cumulative one.



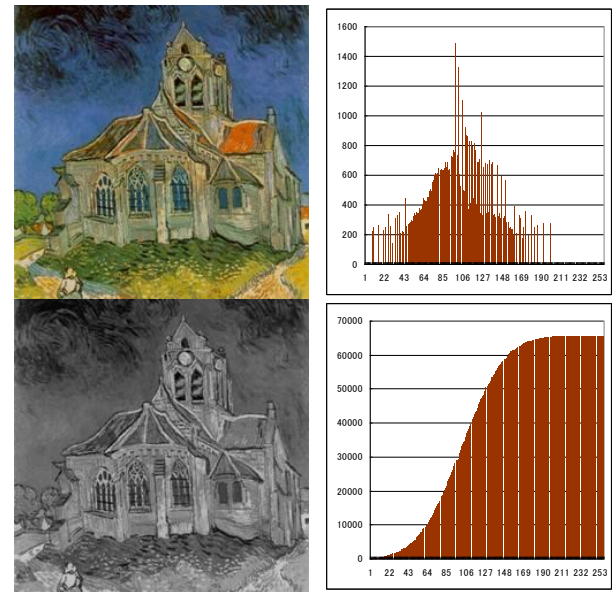
HMGD processed images (color and gray level) and the histogram of the gray level image and its cumulative histogram ($s=5$, estimated STD $\sigma \approx 27.607$).

Fig. 3. Results and the corresponding histograms (I).



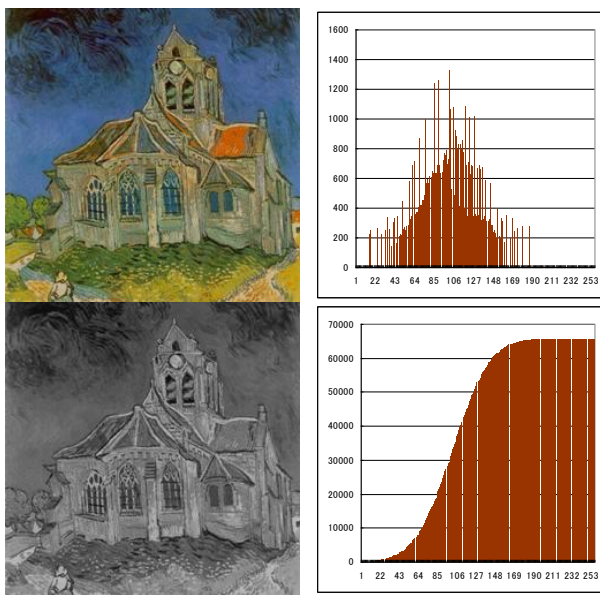
HMGD processed images (color and gray level) and the histogram of the gray level image and its cumulative histogram ($s=9$, estimated STD $\sigma=29.26$).

Fig. 4. Results and the corresponding histograms (II).



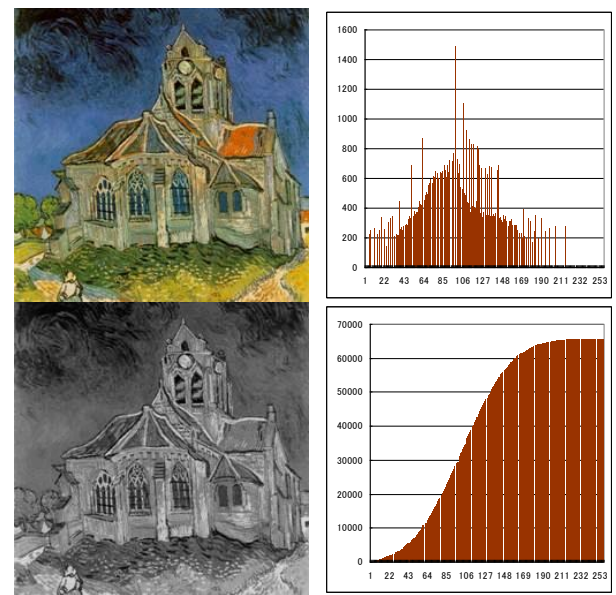
HMGD processed images (color and gray level) and the histogram of the gray level image and its cumulative histogram ($s=17$, estimated STD $\sigma=37.796$).

Fig. 6. Results and the corresponding histograms (IV).



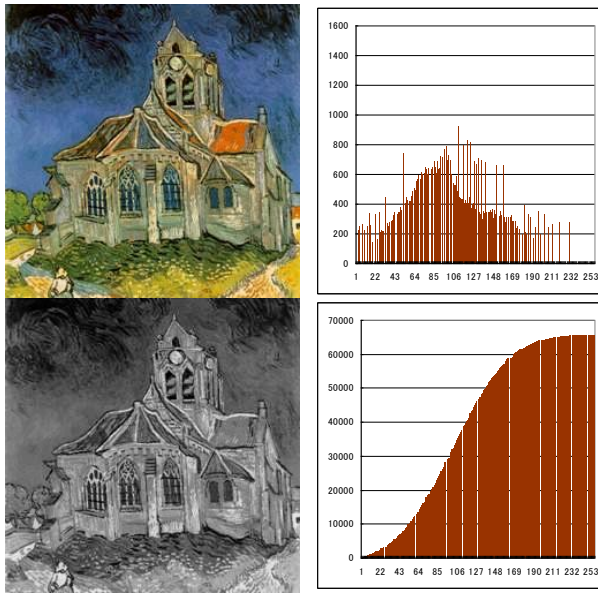
HMGD processed images (color and gray level) and the histogram of the gray level image and its cumulative histogram ($s=13$, estimated STD $\sigma=32.63$).

Fig. 5. Results and the corresponding histograms (III).



HMGD processed images (color and gray level) and the histogram of the gray level image and its cumulative histogram ($s=21$, estimated STD $\sigma=43.596$).

Fig.7. Results and the corresponding histograms (V).



HMGD processed images (color and gray level) and the histogram of the gray level image and its cumulative histogram ($s=25$, estimated STD $\sigma=49.797$).

Fig. 8. Results and the corresponding histograms (VI).

Since the histogram of original gray level image (see Fig.2) roughly shows the single peak and the big mountain shape, the more sampling points (number s) are increased, the more estimated STD σ we have obtained. Then the HMGD processing results in higher contrasted image, i.e. image enhancement.

4. Conclusion

In this paper we have newly proposed the method for the variance estimation using regression analysis, in order to decide the reference histogram in HMGD processing.

As for the method, first, we compute the curvature for the cumulative histogram and detect the brightness peak in the original histogram. Next, we estimate the variance (or standard deviation) around the peak point of the histogram using regression analysis. And we make the reference histogram pattern (Gaussian distribution) based on the estimated variance. Finally, we perform the image transformation by the HMGD processing.

We consider that the experimental results are very good. However, the ability of variance estimation depends on the number of sample points s , that is, the sampling section.

Then we consider that it is better to estimate the variance after the decision of the section, based on the detected peaks and valleys in a histogram.

For further study, we investigate the effectiveness of the proposed variance estimation method when the method is applied to all sections in a histogram where each section contains one peak point between two valleys.

References

1. R. C. Gonzalez and R. E. Woods, *Digital Image Processing* (Addison-Wesley Publishing Company, 1993).
2. B. Jahne, *Digital Image Processing --Concepts, Algorithms, and Scientific Applications-- 4th edition* (Springer, 1995).
3. E. S. Umbaugh, *Computer Vision and Image Processing: A Practical Approach Using CVIP tools* (Prentice Hall PTR, 1998).
4. W. Burger and J. M. Burge, *Principles of Digital Image Processing: Fundamental Techniques* (Springer 2009).
5. T. Izumi, T. Hattori, S. Sugimoto, and T. Takashima, Color Image Arrangement Using Elastic Transform on Principal Component Axis (in Japanese), *Journal of Japan Society of Kansei Engineering* **8**(3) (2009) 667-674.
6. Y. Kawakami, T. Hattori, D. Kutsuna, H. Matsushita, Y. Imai, H. Kawano, R.P.C. Janaka Rajapakse, Automated Color Image Arrangement Method Based on Histogram Matching - Investigation of Kansei impression between HE and HMGD -, *International Journal of Affective Engineering* **14**(2) (2015) ISSN 2187-5413, 85-93.
7. Y. Kawakami, T. Hattori, Y. Imai, H. Matsushita, H. Kawano, R. P. C. Janaka Rajapakse, Kansei Impression and Automated Color Image Arrangement Method, *Journal of Robotics, Networking and Artificial Life* **1**(1) (2014) ISSN 2352-6386, 60-67.
8. Y. Kawakami, T. Hattori, Y. Imai, H. Matsushita, H. Kawano, and R. P. C. Janaka Rajapakse, Automated Color Image Arrangement Method Using Curvature Computation in Histogram Matching, in *Proceedings of International Conference on Artificial Life and Robotics (ICAROB 2015)* (Oita, Japan, 2015) ISBN 978-4-9902880-9-9, 272-277.
9. Y. Kawakami, T. Hattori, Y. Imai, Y. Horikawa, H. Matsushita, R. P. C. Janaka Rajapakse, Automated Processing of Multiple-Brightness Peak Histogram Image Using Curvature and Variance Estimation, *Journal of Robotics, Networking and Artificial Life* **3**(1) (2016) ISSN 2352-6386, 55-60.
10. D. G. Kleinbaum, L. L. Kupper, A. Nizam, E. S. Rosenberg, *Applied Regression Analysis and Other Multivariable Methods* (Brooks/Cole Pub Co, 2013).