J. Appl. Math. & Informatics Vol. 27(2009), No. 1 - 2, pp. 175 - 182
Website: http://www.kcam.biz

APPLICATION OF HISTOGRAM OUTLIER ANALYSIS ON THE IMAGE DEGRADATION MODEL FOR BEST FOCAL POINT SELECTION

HYUNKYUNG SHIN

ABSTRACT. Microscopic imaging system often requires the algorithm to adjust location of camera lenses automatically in machine level. An effort to detect the best focal point is naturally interpreted as a mathematical inverse problem [1]. Following Wiener's point of view [2], we interpret the focus level of images as the quantified factor appeared in image degradation model: $g = f * H + \eta$, a standard mathematical model for understanding signal or image degradation process [3]. In this paper we propose a simple, very fast and robust method to compare the degradation parameters among the multiple images given by introducing outlier analysis of histogram.

AMS Mathematics Subject Classification: 90C90, 90-80, 62M40.
Keywords and phrase : Image degradation model, histogram analysis, best focal point

1. Introduction

In semi-conductor manufacture industry, using advanced technologies of computer vision, visual inspection system for detecting the defects located in the wafers is a vital procedure. One remarkable point to mention in semi-conductor visual inspection is its image resolution level ranges about 0.7 nano-millimeters. To achieve such a microscopic resolution level, it is inevitable to apply huge degree of magnification by making use of series of multiple numbers of the magnifying glasses. In this circumstance, adjustment of focal planes of lenses is a critical concern. Our observation showed that the difference, between the well adjusted- and the less adjusted- focal planes, is revealed in the resulting images in the manner of different degree of corruption: The images taken from the mal-focused lenses are very similar to the images obtained after degradation processing.

Received May 20, 2008. Revised August 7, 2008. Accepted August 20, 2008.
*Corresponding author.

We addressed the following question: "How can we quantify or estimate the focus level of lenses from an image?" This question seemingly can be answered when we consider the corruption process due to mal-focus on the mathematical model of image degradation.

Image restoration has been regarded as a mathematical model equating an inverse process of image degradation and noise. Degradation process usually can be formulated as shown in Eq. (1).

$$g(\vec{x}) = f(\vec{x}) * H(\vec{x}) + \eta(\vec{x}), \tag{1}$$

where g represents degraded image, f source image, H degradation function, η noise, and $\vec{x} \in R^n$ (n indicates the dimension of the image under consideration).

Histogram analysis has played an important role in study on degradation and retrieval of image. In the process of retrieval from degraded image, several forms of statistical discriminant measures are constructed using histogram analysis [4]. A method of co-histogram of two different images is devised and used to retrieve images in [5, 6]. JPEG compression is well known for producing the artifacts due to limitation of Fourier transformation, those artifacts can be considered as a form of degradation, JPEG2000 uses the wavelet transformation ([7]) in order to avoid such problem [8]. In [9], the authors evaluate level of degraded artifacts caused by JPEG compression through comparison to wavelet compressed image. They adopted histogram analysis as main tool for the comparison. In medical imaging area, histogram analysis is also regarded as very useful technique for object segmentation[10].

In Eq. (1), it would be straightforward to estimate σ from image g if we know a reference image f, which is not the case in this study. We try to identify the best focused image among a set of degraded images g's taken the same objects. Our task would be easier if we can assume the noise η has a uniform distribution in Eq. (1). However, we adopt the Gaussian noise model since it is more realistic [2]: $\eta(\vec{x})$ is sampled from the Gaussian distribution and $H(\vec{x}) = e^{\frac{\vec{x}^2}{\sigma^2}}$, where σ stands for the degradation factor. Moreover, our study focus on the case when variance of η is sufficiently large enough not to ignore the noise factor in a degradation model.

This paper is organized as follows: in §2, at first, we study of the effects of degradation parameter and noise on the histogram of image. Then we introduce our method of histogram outlier analysis. In §3, discussions on the study performed in this paper are presented. In §4, various results of using our method on the real samples are displayed.

2. Outlier analysis of histogram

In this section we introduce a mathematical model (M) of defining a quantitative measure to estimate degree of image degradation. This model contains two acting parameters: degree of noise and degradation factor.

$$M(f) \equiv M(f \mid \sigma \; ; \; \eta), \tag{2}$$

where f, σ, and η represent an image, a degradation factor and a noise level, respectively. In §2.1 we provide analytic analysis on the effect of σ on the image histogram. In §2.2 we provide analytic analysis on the effect of η on the image histogram. We will show the mathematical well-definedness of the measure.

2.1 Outlier analysis on effect of H to histogram To see the effect of degradation factor to image histogram we assume no noise, i.e., original image f (analytic) is assumed to be corrupted by degradation H as described in Eq. (3).

$$g(x,y) = f(x,y) * H(x,y), \tag{3}$$

where H is a Gaussian

$$H(x,y) = ce^{-(\frac{x^2+y^2}{\sigma^2})}. \tag{4}$$

Histogram of image g defined in Eq. (1) can be derived explicitly as follows: Suppose f $(f : A \to Z)$ is a bi-level image defined as

$$f(x,y) = \begin{cases} x_0 & \text{where } (x,y) \in A_0 \\ x_1 & \text{where } (x,y) \in A_1 \end{cases}, \tag{5}$$

where $x_0 < x_1$, $A_0 \cup A_1 = A$, and $A_0 \cap A_1 = \emptyset$.

Histogram of f has two peaks at A_0 and A_1 (See Fig. (1)).

$$h_f(n) = \begin{cases} \dfrac{X(A_0)}{X(A)} & \text{where } n = x_0 \\ \dfrac{X(A_1)}{X(A)} & \text{where } n = x_1 \end{cases}, \tag{6}$$

where n denotes gray scale value. Eq. (4) implies $g(x,y)$ is combination of two Gaussian exponential functions.

$$g(x,y) = a_0 e^{-\frac{(x-x_0)^2}{2\sigma^2}} + a_1 e^{-\frac{(x-x_1)^2}{2\sigma^2}}, \tag{7}$$

where $a_0 = X(A_0)$ and $a_1 = X(A_1)$. Estimation of normalized histogram of g is as follows:

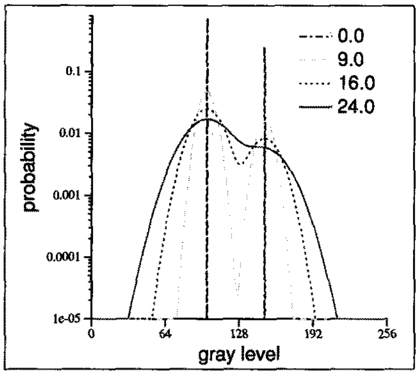
$$h_g(n) = c_0 e^{-\frac{(n-A_0)^2}{2\sigma^2}} + c_1 e^{-\frac{(n-A_1)^2}{2\sigma^2}}, \tag{8}$$

where $c_0 = \dfrac{X(A_0)}{\sqrt{2\pi}\sigma X(A)}$ and $c_1 = \dfrac{X(A_1)}{\sqrt{2\pi}\sigma X(A)}$. Effect of σ to the overall shape of histogram and to the cumulative histogram are presented in Fig. 1.

Right side of Fig. 1 shows the relation as below:

$$h_g|_{\sigma_1}(x) < h_g|_{\sigma_2}(x), \text{ if } \sigma_1 < \sigma_2, \tag{9}$$

where $0 < x < x_0$. This implies that an image with smaller $h_g(x)$ has the better focal position.



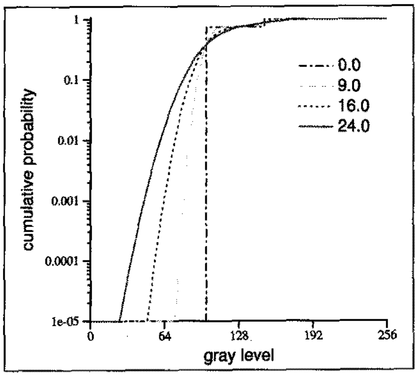


FIGURE 1. Changes of density histogram and cumulative histogram due to degradation factor, σ, are visualized at the left and the right, respectively.

2.2 Outlier analysis on effect of η to histogram In this section our study is focused on effect of presence of heavy noise to image histogram. In case of noise presence, the degradation model Eq. (7) should be transformed as following:

$$g(x,y) = a_0 e^{\frac{(x-x_0)^2}{2\sigma_1^2}} + a_1 e^{\frac{(x-x_1)^2}{2\sigma_2^2}} + a_2 e^{\frac{(x-m)^2}{2\sigma_R}}, \tag{10}$$

where $a_0 = X(A_0)$, $a_1 = X(A_1)$, σ represents degradation factor, a_2, m, and σ_R^2 strength of noise, mean, and variance of random distribution, respectively. To observe the effect of a_2 to histogram, $a_0, a_1, \sigma_1, \sigma_2, \sigma_R$ are assumed to be fixed. The histogram function of g can be established analytically as follows:

$$h_g(n) = c_0 e^{\frac{(n-A_0)^2}{2\sigma_1^2}} + c_1 e^{\frac{(n-A_1)^2}{2\sigma_2^2}} + c_2 e^{\frac{(n-m)^2}{2\sigma_R}}, \tag{11}$$

where $c_0 = \dfrac{X(A_0)(1-c_2)}{\sqrt{2\pi\sigma_1}X(A)}$, $c_1 = \dfrac{X(A_1)(1-c_2)}{\sqrt{2\pi\sigma_2}X(A)}$, $c_2 = a_2$ in Eq. (10),

Fig. 2 shows a parametric relation: $h_g|_{c_2=\alpha_1}(n) < h_g|_{c_2=\alpha_2}(n)$ if $\alpha_1 < \alpha_2$ where $0 < n < x_0$. In order for presenting order relationship, in right side of Fig. 1 cumulative value of h_g at several gray level value.

We have shown that $M(f|\sigma,\eta)$ has actually parametric model property. In §2 we will define a quantitative measure for image degradation due to focal position and will prove well-definedness of our quantifier.

2.3 Outlier analysis We define a cumulative density function (cdf) of histogram as follows:

$$\psi_g(x) = \int_0^x h_g(\tau)d\tau, \tag{12}$$

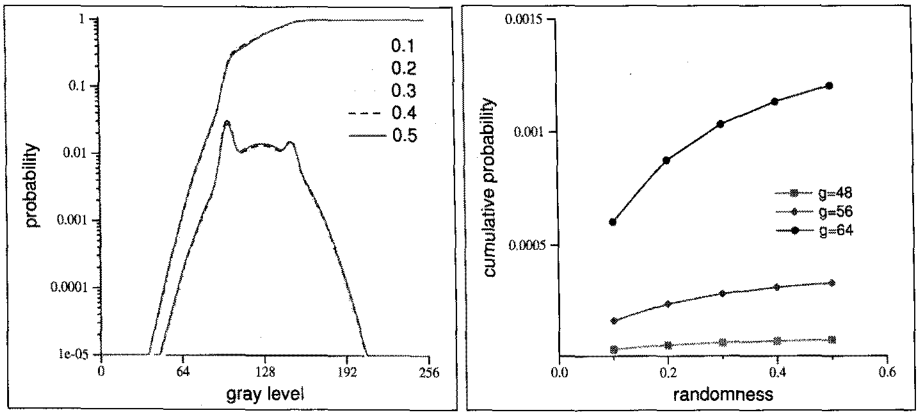


FIGURE 2. At left, changes in pdf and cdf of histograms due to η are visualized. At right, vertical cross sectional values of cdf are presented, where the gray levels are chosen at 48, 56, and 64, respectively.

where h_g is defined in Eq. (11). ψ_g satisfies $0 \leq \psi_g(x) \leq 1$, by definition, and is monotonic since $h_g \geq 0$ for all x, respectively. The latter condition enables us to define an inverse function:

$$\phi_g(y) = \psi_g^{-1}(y), \tag{13}$$

where $0 \leq y \leq 1$.

We can define a new quantifier $Q(g,y) : \iota \times R \to R$, where ι denotes a space of images. Set

$$Q(g,y) \equiv \phi_g(y). \tag{14}$$

As we introduce a new quantifier for the first time, we present a brief description as follows:

well-definedness: For any image $g \in \iota$, a histogram of g is well-defined as shown in Eq. (11).This implies that $Q(g,y)$ is well-defined for any image.

uniquess of mapping: It is obvious $Q(g,y) = Q(g',y)$ if $g \equiv g'$, which guarantees the uniqueness of Q.

stability of mapping: Continuity of h_g implies existence of a positive real number M satisfying $\|Q(g_1,y) - Q(g_2,y)\|_p \leq M\|g_1 - g_2\|_p$ for any $1 \leq p \leq \infty$.

Set

$$D_\epsilon = \Big\{ x \mid 0 < x < \epsilon \cup 1-\epsilon < x < 1 \Big\}, \tag{15}$$

where $0 < \epsilon < 1$. Intuitively, outlier analysis of histogram implies that estimation of $Q(g,y)$ where $y \in D_\epsilon$ for small positive value of ϵ.

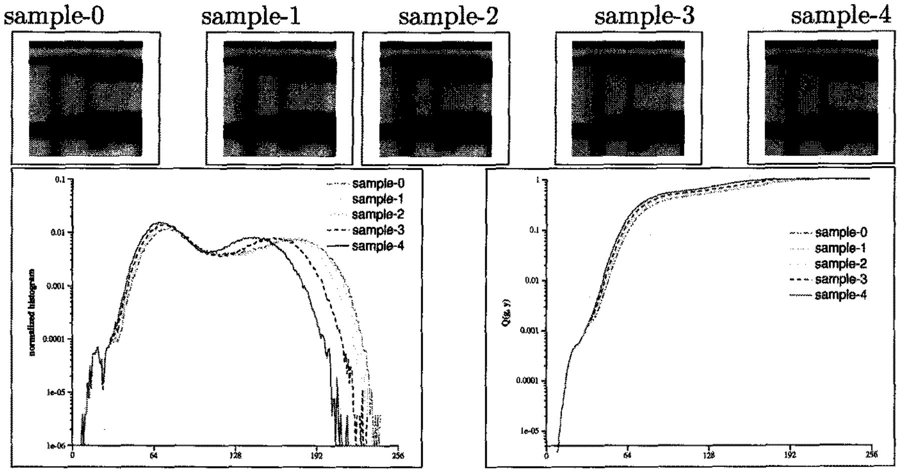


FIGURE 3. Left panel illustrates the plots of the normalized histogram of sample images. Right panel illustrates the computed values of $Q(g, y)$. This results show the sample-4 is the one with the best focal point.

The result shown in Fig. 1, 2 guarantees that we can find $\epsilon > 0$ such that if $Q(g_1, y) \leq Q(g_2, y)$ for $y \in D_\epsilon$ then g_2 is the one with smaller value of σ, i.e., better focused.

3 Discussion

Our approach proposed in this paper is computationally very efficient as the following. Creation of the histogram of an image requires $N \times M$ reading and addition instructions where N and M represents width and height in pixel of the image, construction of the *PDF* requires 256 addition and multiplication, and *CDF* construction requires 256 addition operations. This makes total of exact $2 \times N \times M + 3 \times 256$ instructions for each image.

Typically, quantification of image degradation can be achieved better in quality performance using frequency domain methods such as FFT and wavelet. This is caused by the fact that convolution factor in degradation model like (Eq. (1)) can be decoupled out easily from the frequency domain transformation. However, in the restrictive situation such as a real time process, which requires speed rather than quality, those frequency domain methods tending to take much more processing time are useless.

4. Experimental results

The method developed in this paper is applied to the real world sample images. The results are presented in Fig. 3–4.

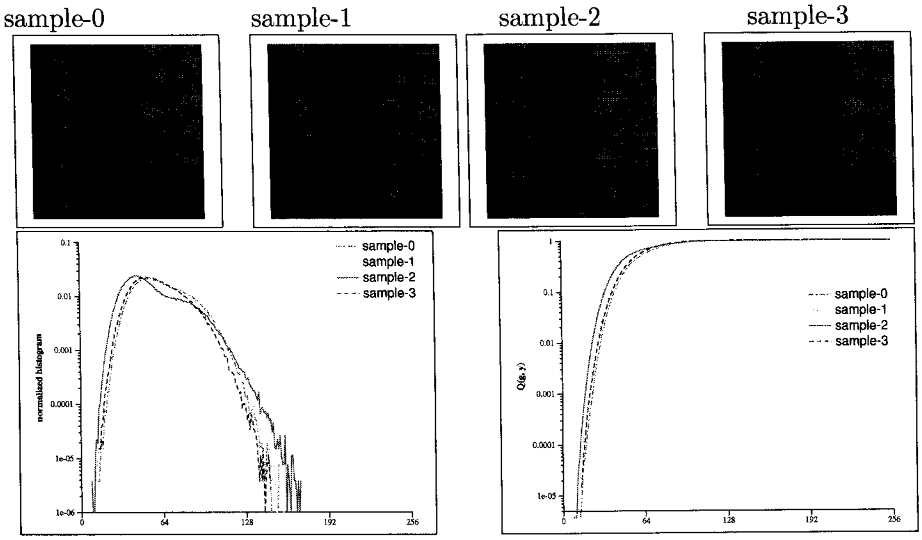


FIGURE 4. Left panel illustrates the plots of the normalized histogram of sample images. Right panel illustrates the computed values of $Q(g, y)$. This results show the sample-2 is the one with the best focal point.

The five images used in Fig. 3 have different focal points. As seen in the top of the figure, the images are severely damaged by degradation and heavy noise. In human eyes, it is not clear which sample is the best focused. Our result using $Q(g, y)$ reveals that sample-4 is the one best focused. The four sample images appeared in Fig. 4 are also severely damaged. In human eyes, sample-2 is the one best focused. Our quantifier $Q(g, y)$ confirms it.

Acknowledgement

We appreciate to KLA-Tencor for approving us the sample images used in Fig. 3–4.

REFERENCES

1. Tarantola, A., *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia, PA, 2005.
2. Wiener, N., *Cybernetics or Control and Communication in the Animal and the Machine*, The MIT Press, Cambridge, MA, 1961.
3. Gonzalez, R.C., and Woods, R.E., *Digital Image Processing*, Prentice Hall, Upper Saddle River, NJ, 2002.
4. Brunelli, B., and Mich, O., *Histograms Analysis for Image Retrieval*, ITC-irst, Interactive Sensory Systems Division, I-38050 Povo, Trento, Italy, Feb., 2000.

5. Hao, P., Zhang, C., and Dang, A., *Co-Histogram and Image Degradation Evaluation*, International Conference on Image Analysis and Recognition (ICIAR), pp. 195–203, Sep., 2004.
6. Hao, P., and Chen, Y., *Co-Histogram and Its Application in Video Analysis*, In Proceedings of IEEE International Conference on Multimedia and Expo (ICME), Taiwan, 2004.
7. Taubman, D.S., Marcellin, M.W., *JPEG2000: standard for interactive imaging*, Proceedings of the IEEE, Vol. 90, No. 8, pp. 1336-1357, 2002.
8. Datcu, M., Schwarz, G., Schmidt, K., and Reck, C., *Histogram analysis of JPEG compressed images as an aid in image deblocking*, DCC '95: Proceedings of the Conference on Data Compression, IEEE Computer Society, pp. 425, Washington, DC, 1995.
9. Kalyanpur, A., Neklesa, V.P., Taylor, C.R., Daftary, A.R., and Brink, J.A., *Evaluation of JPEG and Wavelet Compression of Body CT Images for Direct Digital Teleradiologic Transmission*, Radiology, v. 217, pp. 772–779, 2000.
10. Mendonc, P.R.S., Padfield, D.R., Ross, J.C., Miller, J.V., Dutta, S. and Gautham, S.M., *Quantification of Emphysema Severity by Histogram Analysis of CT Scans*, Medical Image Computing and Computer-Assisted Intervention MICCAI, Springer-Verlag, pp. 738–744, 2005

Hyunkyung Shin received her Ph. D in applied mathematics and statistics from State University of New York at stonybrook. Since 2008 she joined in department of mathematics and information of Kyung Won University as assistant professor. Her research interests are theory of computation, image processing, computer vision, and machine learning based on neural networks.

Department of Mathematics & Information, Kyungwon University, Seongnam, 461–701, Korea.

e-mail: `hyunkyung@kyungwon.ac.kr`