A Revisit of Various Thresholding Methods for Segmenting the Contact Area Images

Yang Xu[a1]

[a] School of Mechanical Engineering, Hefei University of Technology, Hefei, 230009, China

1 Image Segmentation - Fundamentals

Image segmentation is served as the pre-processing step for machine vision (Haralick & Shapiro, 1991). Among all different types of image segmentation methods, *thresholding-based* segmentation is one of the simplest and well-developed approach. Thresholding-based segmentation (or *thresholding* for short) is a method that manually/automatically find the optimal thresholds (gray levels) to divide all pixels of a gray-level image into N groups and assign obviously different gray values to all pixels in the same group. When $N > 2$, it is referred to as *multilevel* thresholding. When $N = 2$, the multilevel thresholding degrades to the *bilevel* thresholding.

In the contact mechanics experiments, optimal method is mainly used to identify the real contact area between two solid bodies with at least one being transparent. The contact spots and its vicinity under appropriate illumination show strong contrast of light intensity inside and outside the contact area. Therefore, it is nature to apply bilevel thresholding to distinguish the contact area (object) from the non-contact area (background). Bilevel thresholding assigns obviously different gray values to contact and non-contact area. The bilevel thresholding process is known as *binarization*.

Bilevel thresholding methods are divided into two groups: *global* thresholding and *local* thresholding. The former one results in a constant threshold (gray level). The latter one has a spatially varied thresholds which are determined at local regions. In the present study, several representative bilevel global thresholding methods are applied to

[1]Corresponding author: Yang.Xu@hfut.edu.cn

binarize the gray-level images captured of the nominal contact area which is composed of contact and non-contact areas.

Consider a l-bit gray-level digital image with N_x and N_y pixels in the two orthogonal directions. The gray-level g is an integer varies between 0 and $2^l - 1$. Let (i, j) be the integer indices for locating pixels in the digital image. Define a mapping function $f : N_x \times N_y \to g$. The gray level of the pixel at location (i, j) can be denoted as $f(i, j)$. Before we introduce each bilevel image segmentation methods, let us first get familiar with the commonly adopted concepts and parameters.

The two distinct regions segmented by the bilevel thresholding method is commonly referred to as *object* and *background*. In the present study, we will replace them with *contact area* and *non-contact area*. Contact area represents a group of pixels whose gray value is less than or equal to the threshold (k) and is denoted by C_1. Similarly, C_2 represents the remaining pixels whose gray value is larger than the threshold:

$$C_1 = \{(i, j)|f(i, j) \leq k\}, \quad C_2 = \{(i, j)|f(i, j) > k\}.$$

The threshold k is automatically found by the bilevel global thresholding method based on a certain criterion. C_1 and C_2 in the binarized image are assign 0 and $2^l - 1$.

Generally, all bilevel global thresholding method relies on the histogram of the gray level image. Let the probability of occurrence of gray level g be $p(g)$ so that $\sum_{g=0}^{2^l-1} p(g) = 1$. We can easily calculating $p(g)$ by counting the number of pixels with gray level of g. Using the histogram $p(g)$, we can further define several important statistical parameters:

- Moments of i^{th} order m_i (Tsai, 1985)

$$m_i = \frac{1}{N_x N_y} \sum_{i=0}^{N_x} \sum_{j=0}^{N_y} f(i, j)^i \tag{1}$$

$$= \sum_{g=0}^{2^l-1} g^i p(g), \quad i = 0, 1, 2, 3, \cdots \tag{2}$$

where $m_0 = 1$ represents the probability conservation.

- Mean grey value μ and standard deviation of grey value σ

$$\bar{g} = m_1 = \sum_{g=0}^{2^l-1} g p(g), \tag{3}$$

$$\sigma = \sqrt{m_2} = \sqrt{\sum_{g=0}^{2^l-1} g^2 p(g)}. \tag{4}$$

- Occupance percentage of C_1 and C_2: A^* and $1 - A^*$

$$A^*(k) = \sum_{g=0}^{k} p(g), \quad 1 - A^*(k) = \sum_{g=k+1}^{2^l-1} p(g). \tag{5}$$

Due to the truncation error when calculating $p(g)$, $A^*(k)$ calculated by the above equation may result in a value larger than 1. Therefore, a more reliable way of calculating A^* is to count number of pixels within C_1:

$$A^* = \frac{N_{g \le k}}{N_x N_y}. \tag{6}$$

where $N_{g \le k}$ represents the number of pixels whose gray level $g \le k$.

- Mean gray values ($\bar{g}_1$ and $\bar{g}_2$) inside C_1 and C_2

$$\bar{g}_1(k) = \sum_{g=0}^{k} g p(g) / A^*(k), \quad \bar{g}_2(k) = \sum_{g=k+1}^{2^l-1} g p(g) / \left[1 - A^*(k)\right]. \tag{7}$$

- Standard deviation of gray values (σ_1 and σ_2) inside C_1 and C_2

$$\sigma_1^2(k) = \sum_{g=0}^{k} (g - \bar{g}_1)^2 p(g) / A^*(k), \tag{8}$$

$$\sigma_2^2(k) = \sum_{g=k+1}^{2^l-1} (g - \bar{g}_2)^2 p(g) / \left[1 - A^*(k)\right]. \tag{9}$$

2 Bilevel global thresholding methods

- Otsu's method (Otsu, 1979) [1]

  Python function: `Otsu1979()`.

- Moment-preserving method (Tsai, 1985) [2]

  Python function: `Tsai1985()`.

- Iterative method (Ridler & Calvard, 1978) [3]

  Python function: `RC1978()`.

- K-means clustering [4]

  Python function: `K_means()`.

- Max entropy method (Kapur et al., 1985) [5]

  Python function: `KSW1985()`.

- Fuzzy method (Huang and Wang, 1995) [6]

  Python function: `HW1995()`.

- Minimum error (Kittler & Illingworth, 1986; Sahli et al., 2018) [7, 8]

  Python function: `KI1986()`, `Sahli_etal_2018()`.

2.1 Otsu's method

The well-known Otsu's method [1] relies on the minimization of *within-class variance* $\sigma_w^2(k')$:

$$k = \arg\min \sigma_w^2(k'), \tag{10}$$

where

$$\sigma_w^2(k') = A^*(k')\sigma_1^2(k') + [1 - A^*(k')]\, \sigma_2^2(k'). \tag{11}$$

Greedy method can be applied to calculate $\sigma_w^2(k')$ for all possible $k' = 0, 1, \cdots, 2^l - 1$ and pick $k = k'$ associated with the minimum $\sigma_w^2(k')$.

2.2 Moment-preserving method

Tsai [2] proposed a moment-preserving criterion to effectively determine the threshold. Let all pixels within C_1 and C_2 of binarized images being assigned the gray level of g_0 and g_1. Notice that g_0 and g_1 are not necessarily to be 0 and $2^l - 1$. According to Eq. (2) the moments of the binarized image are

$$m_i^{'} = A^* g_0^i + [1 - A^*]\, g_1^i. \tag{12}$$

Moment-preserving criterion requires the moments (with $i = 1, 2, 3$) of the original and binarized images should remain the same, then we have the following three equations

$$A^* g_0 + [1 - A^*]\, g_1 = m_1, \tag{13}$$

$$A^* g_0^2 + [1 - A^*]\, g_1^2 = m_2, \tag{14}$$

$$A^* g_0^3 + [1 - A^*]\, g_1^3 = m_3, \tag{15}$$

with three unknowns A^*, g_0 and g_1. An explicit formulation of A^* was given by Tsai in Appendix A.1. of Ref. [2]:

$$\begin{aligned}
c_d &= \begin{vmatrix} m_0 & m_1 \\ m_1 & m_2 \end{vmatrix}; \quad c_0 = (1/c_d) \begin{vmatrix} -m_2 & m_1 \\ -m_3 & m_2 \end{vmatrix}; \quad c_1 = (1/c_d) \begin{vmatrix} m_0 & -m_2 \\ m_1 & -m_3 \end{vmatrix}; \\
z_0 &= \frac{1}{2}\left(-c_1 - \sqrt{c_1^2 - 4c_0}\right); \quad z_1 = \frac{1}{2}\left(-c_1 + \sqrt{c_1^2 - 4c_0}\right); \\
A^* &= \begin{vmatrix} 1 & 1 \\ z_0 & z_1 \end{vmatrix}.
\end{aligned} \tag{16}$$

where $|\bullet|$ is the determinant symbol.

Now, the problem degrades to a p-tile problem where the threshold results in $A^* = \sum_{g=0}^{k} p(g)$. Since k is an integer, the aforementioned identity cannot always be satisfied, the following approximation is used

$$k = \arg\min |A^* - \sum_{g=0}^{k'} p(g)|. \tag{17}$$

2.3 Iterative method

Iterative method (also known as isodata algorithm) was proposed by Ridler and Calvard [3] to iteratively correct the threshold as the mean of the mean gray values of C_1 and C_2. At i^{th} iteration, the threshold before iteration is k_{i-1}. The mean gray values $\bar{g}_1(k_{i-1})$ and $\bar{g}_2(k_{i-1})$ is (see Eq. (7))

$$\bar{g}_1(k_{i-1}) = \sum_{g=0}^{k_{i-1}} gp(g)/A^*(k_{i-1}), \quad \bar{g}_2(k_{i-1}) = \sum_{g=k_{i-1}+1}^{2^l-1} gp(g)/[1 - A^*(k_{i-1})]\,. \tag{18}$$

where $A^*(k_{i-1}) = \sum_{g=0}^{k_{i-1}} p(g)$. The threshold k_i after iteration

$$k_i = \frac{1}{2}\left[\bar{g}_1(k_{i-1}) + \bar{g}_2(k_{i-1})\right]. \tag{19}$$

The iteration stops until the threshold converges: $k_i = k_{i-1} = k$. $k_0 = (\min(f(i,j)) + \max(f(i,j)))/2$.

2.4 K-means clustering (see standard textbook of Machine Learning)

K-means clustering is a method of partitioning all pixels of gray-level image into K clusters in which the gray values of all pixels within each cluster has a minimum deviation from its mean gray value. Strictly speaking, it is not a thresholding-based segmentation method since K clusters is not obtained by simply thresholding the images with $K - 1$ thresholds. Besides its application in the image segmentation, it is also widely used in machine learning and data science. Authors cannot find the very first paper in which K-means clustering was applied to the bilevel thresholding. Find the textbook of machine learning which contains K-means clustering.

For bilevel thresholding, K-means clustering partitions all pixels' gray values into two groups, C_1 and C_2. The corresponding mean gray value of each group is $\bar{g}_1$ and $\bar{g}_2$, respectively. The partition rule is to check the absolute difference between a pixel's gray value g and $\bar{g}_i$, $i = 1, 2$. If $|g - \bar{g}_1| \leq |g - \bar{g}_2|$, the corresponding pixels belong to C_1 and vice versa. Since $\bar{g}_1$ and $\bar{g}_2$ are unknown before the initial iteration, we assume $\bar{g}_1 = \min(f(i,j))$ and $\bar{g}_2 = \max(f(i,j))$. After the partition, the mean gray values $\bar{g}_i$, $i = 1, 2$, will be updated. The partition will be repeated until the mean gray values $\bar{g}_i$, $i = 1, 2$, remain the same.

2.5 Max entropy method

A group of algorithms relies on the maximizing entropy to obtain the optimal threshold. Entropy is a concept in the information theory which was brought to the image segmentation by Pun [9] for the first time. Kapur et al. [5] proposed the following entropies for C_1 and C_2

$$H_1(k) = -\sum_{g=0}^{k} \frac{p(g)}{A^*(k)} \ln \frac{p(g)}{A^*(k)}, \tag{20}$$

$$H_2(k) = -\sum_{g=k+1}^{2^l-1} \frac{p(g)}{1 - A^*(k)} \ln \frac{p(g)}{1 - A^*(k)}. \tag{21}$$

If $A^*(k) = 0$ or 1, its contribution to $H_1(k)$ or $H_2(k)$ will be ignored. The threshold is the argument of $H(k) = H_1(k) + H_2(k)$ when it takes the maximum value:

$$k = \arg\max H(k'). \tag{22}$$

2.6 Fuzzy method

Huang and Wang [6] proposed the following membership function of the gray value $g \in [0, 2^l - 1]$ and the threshold k

$$\mu(g,k) \begin{cases} = \dfrac{1}{1 + |g - \bar{g}_1|/C} & \text{if} \quad g \leq k, \\ = \dfrac{1}{1 + |g - \bar{g}_2|/C} & \text{if} \quad g > k, \end{cases} \tag{23}$$

where $\bar{g}_1$ and $\bar{g}_2$ are calculated by Eq. (7). C must be selected to strictly guarantee $\mu(g,k) \in [1/2, 1]$. A reasonable range for C is

$$C \geq \max(\bar{g}_1, |k - \bar{g}_1|, |2^l - 1 - \bar{g}_2|, |k - \bar{g}_2|). \tag{24}$$

In the present study, C is taken the maximum value of the right hand side of Eq. (24) for all $k \in [0, 2^l - 1]$. The membership function quantifies the relationship of an individual pixel with either contact or non-contact area in a fuzzy way. $\mu \to 1$ means the associated pixel is more likely to be included in the present region.

Huang and Wang defined an alternative entropy of a gray-level image

$$E(k) = \frac{1}{n_x n_y \ln 2} \sum_{g=0}^{2^l-1} S(\mu(g,k)) p(g), \tag{25}$$

where $S(\mu)$ is Shannon's function:

$$S(\mu) = -\mu \ln(\mu) - (1 - \mu) \ln(1 - \mu).$$

If $\mu \to 1$, the limiting value of S is zero. The threshold is the argument of the minimum of the entropy:

$$k = \arg\min E(k'). \tag{26}$$

Notice that an improved threshold with better accuracy may be obtained with an empirical correction. This conflicts with our idea of a parameter-free algorithm.

2.7 Minimum error method - Kittler and Illingworth (1985)

Kittler and Illingworth assumed that the gray-level image is bimodal so that the histogram $p(g)$ has two distinct peaks at lower and higher range of gray values. Bimodal assumption is commonly adopted in the early development of the bilevel global thresholding methods. Those methods relies on the locations of two peaks in the histogram to find the valley. The corresponding gray value of the valley is the threshold. This bimodal shape-dependent methods may not work for those images with no clear bimodal feature.

Let $p(g|C_1)$ and $p(g|C_2)$ be the posterior conditional probability of gray value given corresponding pixel in C_1 and C_2, respectively. Kittler and Illingworth [7] used the following Gaussian distribution to approximate $p(g|C_i)$ where $i = 1, 2$:

$$p(g|C_i, k) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(g - \bar{g}_i)^2}{2\sigma_i^2}\right], \tag{27}$$

where $\bar{g}_i$ and σ_i ($i = 1, 2$) are all unknowns. Therefore, the histogram $p(g)$ can be explicitly expressed as

$$p(g) = A^*(k)p(g|C_1) + [1 - A^*(k)]p(g|C_2). \tag{28}$$

Fitting the right hand side of Eq. (28) to the histogram of the image, we can get the values of A^*, σ_1, σ_2, $\bar{g}_1$ and $\bar{g}_2$. The optimum value of k is associated with the intersection of $A^*P(g|C_1)$ and $(1 - A^*)P(g|C_2)$.

2.8 Minimum error method - Sahli et al. (2018)

Sahli et al. [8] (see the supporting information in Ref. [8] developed a slightly different minimum error method. Eq. (28) is still used to approximate the histogram $p(g)$. The conditional probability of gray value within $C1$ (contact area) is still approximated by the Gaussian distribution

$$p(g|C_1) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(g-\bar{g}_1)^2}{2\sigma_1^2}\right], \tag{29}$$

where σ_1 and $\bar{g}_1$ are unknowns. The assumed form of $p(g|C_2)$ associated with C_2 (out-of-contact area) is no longer Gaussian

$$p(g|C_2) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(g-\bar{g}_2)^2}{2\sigma_2^2}\right] \log\left\{1+\exp\left[0.1(g-d)\right]\right\}. \tag{30}$$

Then, Eq. (28) for histogram $p(g)$ can be rewritten as

$$p(g) = A^* p(g|C_1) + (1-A^*) p(g|C_2). \tag{31}$$

After fitting the right hand side of Eq. (31) with 6 unknowns (namely, A^*, σ_1, σ_2, $\bar{g}_1$, $\bar{g}_2$, d) to the histogram of the image, the optimized threshold k is associated with the intersection of $A^* p(g|C_1)$ and $(1-A^*)p(g|C_2)$.

References

[1] Otsu N. A threshold selection method from gray-level histograms[J]. IEEE transactions on systems, man, and cybernetics, 1979, 9(1): 62–66.

[2] Tsai W-H. Moment-preserving thresolding: A new approach[J]. Computer vision, graphics, and image processing, 1985, 29(3): 377–393.

[3] Ridler T, Calvard S, others. Picture thresholding using an iterative selection method[J]. IEEE Trans. Syst. Man Cybern, 1978, 8(8): 630–632.

[4] Hartigan J A. Clustering algorithms[M]. [S.l.]: John Wiley & Sons, Inc., 1975.

[5] Kapur J N, Sahoo P K, Wong A K. A new method for gray-level picture thresholding using the entropy of the histogram[J]. Computer vision, graphics, and image processing, 1985, 29(3): 273–285.

[6] Huang L-K, Wang M-J J. Image thresholding by minimizing the measures of fuzziness[J]. Pattern recognition, 1995, 28(1): 41–51.

[7] Kittler J, Illingworth J. Minimum error thresholding[J]. Pattern recognition, 1986, 19(1): 41–47.

[8] Sahli R, Pallares G, Ducottet C, et al. Evolution of real contact area under shear and the value of static friction of soft materials[J]. Proceedings of the National Academy of Sciences, 2018, 115(3): 471−476.

[9] Pun T. A new method for grey-level picture thresholding using the entropy of the histogram[J]. Signal processing, 1980, 2(3): 223−237.