The Vagueness Measure: A New Interpretation and an Application to Image Thresholding

József Dombi* and Gergely Gulyás†

Abstract

Here, we introduce a new interpretation of the vagueness measure (which appeared in an earlier work) and an application for this approach. If the vagueness measure is computed for the distribution function of a given population, the value obtained gives a similar characteristic as the standard deviation of the population. Based on this property, a new global thresholding algorithm was developed that generalizes the idea of Otsu's optimality criterion by the means of continuous-valued logic. The performance of this method is compared with other commonly used algorithms to validate the usefulness of the proposed approach. Although the purpose of this algorithm is to threshold a grayscale image (which can be a useful step in the segmentation process of biological and medical images), it can be generalized for other tasks that require the separation of two or more populations, characterized by real values.

Keywords: Image segmentation; Global thresholding; Fuzziness measure; Pliant system; Vagueness functions; Vagueness measure; Ignorance Functions; Weak Ignorance Functions

1 Introduction

In the field of fuzzy sets, the fuzziness measures are used to reflect the uncertainty of the membership functions. The vagueness measure [9] was derived from fuzziness measures and it is a part of the Pliant system [11] which is a coherent continuous-valued logical system. A vagueness measure can be constructed by means of vagueness functions, which are closely related to entropy functions like Shannon's entropy function (as we showed earlier). In this paper we present a new interpretation of the vagueness measure. When applied to a distribution function, the result value gives a similar characteristic as the standard deviation of the given population.

In the field of image processing, segmentation is one of the most important tasks. The goal of segmentation is to divide the image into disjoint regions so that these

---

*6720, Szeged, Árpád tér 2, Hungary, University of Szeged, E-mail: `dombi@inf.u-szeged.hu`
†University of Szeged, E-mail: `gulyasg@inf.u-szeged.hu`

classes represent different objects. A frequently used technique for segmentation is the global thresholding, whose method is very useful in biological and medical image processing. Generally, the task requires to choose an appropriate threshold value for a given property and this value divides the pixels of the image into two classes based on the property: the pixels whose property value are below of the threshold constitute the first class, while the others belong to the second class. In many cases the property is the grayscale value of the pixels, but it can be other feature value which were computed for each pixels. Many algorithms have been introduced in the literature that can choose a good threshold value [15, 18, 23, 24, 26, 29, 33]. Some of these methods are based on fuzzy set theory and are able to handle any ambiguity of data [3, 4, 6, 7, 13, 14, 27, 28].

In this paper we will present the vagueness measure in a new context because we apply it on distribution functions. We will show that this interpretation is related to the standard deviation. Based on this property, we give a new application of the vagueness measure in a global thresholding algorithm which generalizes the idea of Otsu's method. The proposed approach will be compared with other well-known algorithms in order to show the effectiveness of the algorithm.

2 Preliminaries

Let f be a 2-dimensional discrete image function which is defined in such a way that

$$f \colon \{0, \ldots, M-1\} \times \{0, \ldots, N-1\} \to \{0, \ldots, L-1\},$$

where L denotes the maximum intensity level. Let h denote the histogram (function) of the image which gives the number of occurrences at a given gray level:

$$h(q) = |\{(x,y) : f(x,y) = q \text{ and } (x,y) \text{ is a pixel}\}|,$$

where $|.|$ is the cardinality of a given set. The normalized histogram p gives the probability of occurrence of a given gray level:

$$p(q) = P(\text{intensity} = q) = \frac{h(q)}{\sum_{i=0}^{L-1} h(i)}$$

For a given threshold value t, we can define the a priori probabilities of the background and object:

$$p_B^{(t)} = \sum_{q=0}^{t} p(q) \qquad p_O^{(t)} = \sum_{q=t+1}^{L-1} p(q)$$

As we mentioned earlier, which part of the intensities belongs to the object depends on the image, but for simplicity we will take the object to be the brighter part of the image. Let c denote the cumulative distribution function such that

$$c(q) = P(\text{intensity} \leq q) = \sum_{i=0}^{q} p(i).$$

Consider the cumulative distribution functions of the background and object for a given threshold t:

$$c_B^{(t)}(q) = \frac{\sum_{i=0}^{q} h(i)}{\sum_{i=0}^{t} h(i)} \qquad\qquad q = 0, \ldots, t$$

$$c_O^{(t)}(q) = \frac{\sum_{i=t+1}^{q} h(i)}{\sum_{i=t+1}^{L-1} h(i)} \qquad\qquad q = t+1, \ldots, L-1$$

We can assign a binary image to a given grayscale image f and a given threshold intensity t in the following way:

$$f_t(x, y) = \begin{cases} 1 & \text{if } f(x, y) \leq t \\ 0 & \text{otherwise} \end{cases}$$

In general, a *global thresholding algorithm* is one that determines a single threshold value and thresholds the entire image with that value.

3 Vagueness functions and vagueness measure

In Fuzzy Theory [31, 32] the membership function (denoted by μ) can be regarded as the approximation of the characteristic function belonging to crisp sets. Then, a fuzziness measure expresses the distance between the characteristic function and a given membership function, and reflects the uncertainty of the membership function. In the following sections we introduce the vagueness functions and the vagueness measure [9] which belong to the Pliant system (which will be also introduced). The vagueness measure can be interpreted as a fuzziness measure in the Pliant system but we give a new interpretation of the vagueness measure on distribution functions (when it does not play the role of a fuzziness measure).

3.1 Pliant system

The Pliant system [8, 10, 11] is a subclass of continuous valued logic where the fuzzy logical operators like negation, conjunction and disjunction, and the fuzziness measure constitute a coherent system. In the Pliant logic each operator is defined by one generator function. In fuzzy logic, several different generator functions are used. One of the main features of the Pliant concept that consists of infinitely many negations [8]. In the following we shortly introduce the definition of the Pliant system and the necessary notions from fuzzy logic.

Definition 1. *We say that $n(x)$ is a negation if $n\colon [0,1] \to [0,1]$ satisfies the following conditions:*

C1: $n : [0,1] \to [0,1]$ *is continuous*   *(Continuity)*
C2: $n(0) = 1$, $n(1) = 0$ *(Boundary conditions)*
C3: $n(x) < n(y)$ *for $x > y$*   *(Monotonicity)*
C4: $n(n(x)) = x$ *(Involution)*

Using the general representation theorem we have for the strict t-norm (conjunctive operator) and strict t-conorm (disjunctive operator) the following definitions.

Definition 2. *Let $f_c(x) : [0,1] \to [0,\infty]$ be continuous and strictly decreasing monotone generator function. Then*

$$c(x,y) = f_c^{-1}(f_c(x) + f_c(y)). \tag{1}$$

is a strict t-norm (conjunctive operator).

Definition 3. *Let $f_d(x) : [0,1] \to [0,\infty]$ be continuous and strictly increasing monotone generator function. Then*

$$d(x,y) = f_d^{-1}(f_d(x) + f_d(y)). \tag{2}$$

is a strict t-conorm (disjunctive operator).

Note: In Pliant logic, c stands for the conjunctive operator and d stands for the disjunctive operator. Those familiar with fuzzy logic theory will find that the terminology used here is slightly different from that used in standard texts [1, 2, 5, 12, 17, 21].

Definition 4. *If $f_c(x)$ and $f_d(x)$ are related in an inverse manner, i.e.*

$$f_c(x)f_d(x) = 1, \tag{3}$$

then we will call the generated connectives a Pliant system. It has been shown that only this system of strict t-norms and t-conorms is equipped with infinitely many negations [8].

Definition 5. *The general form of the multiplicative Pliant system is*

$$o_\alpha(x,y) = f^{-1}\left((f^\alpha(x) + f^\alpha(y))^{1/\alpha} \right) \tag{4}$$

where $f : [0,1] \to [0,\infty]$ is a continuous and strictly decreasing function which is the generator function of the strict t-norm operator.

If $\alpha > 0$,   *than* $o_\alpha(x,y)$ *conjunctive operator (t-norm).*
If $\alpha < 0$,   *than* $o_\alpha(x,y)$ *disjunctive operator (t-conorm).*   (5)

The corresponding negation is:

$$\eta_\nu(x) = f^{-1}\left(f(\nu_0)\frac{f(\nu)}{f(x)} \right) \qquad or \qquad \eta_{\nu_*}(x) = f^{-1}\left(\frac{f^2(\nu_*)}{f(x)} \right) \tag{6}$$

We have to note that η_ν and η_{ν_*} are well defined and satisfy the conditions of Definition 1 (see in [8]), thus the negations can be interpreted when $f(x) = 0$. The parameter ν, like ν_* is the neutral value of the negation and can be interpreted as the strictness of the negation.

Because the generator function is determined up to a multiplicative constant, we can arrange it such that $f(\nu_0) = 1$ and so $\eta_\nu(x) = f^{-1}(f(\nu)/f(x))$. If $f(\nu_0) = f(\nu) = 1$, then we get the following.

Definition 6. *Let* $f(x) : [0, 1] \rightarrow [0, \infty]$ *be continuous and strictly decreasing monotone function. Then*

$$\eta(x) = f^{-1}\left(\frac{1}{f(x)}\right) \tag{7}$$

is the standard Pliant negation function.

In the Pliant system the conjunctive and disjunctive operators exist in weighted forms which are defined by the following formulas:

$$c(\underline{\mathbf{w}}, \underline{\mathbf{x}}) = f^{-1}\left(\sum_{i=1}^{n} w_i f(x_i)\right) \quad \text{and} \quad d(\underline{\mathbf{w}}, \underline{\mathbf{x}}) = f^{-1}\left(\frac{1}{\sum_{i=1}^{n} \frac{w_i}{f(x_i)}}\right). \tag{8}$$

Thus, we can define a bivariate conjunction operator where the weights are equal, and what we call mean conjunction operator.

Definition 7. *Let* $\bar{c}(x, y)$ *denote the mean conjunction operator in bivariate case:*

$$\bar{c}(x, y) = f^{-1}\left(\frac{1}{2}\left(f(x) + f(y)\right)\right) \tag{9}$$

where $f : [0, 1] \rightarrow [0, \infty)$ *is a continuous, strictly decreasing generator function such that* $f(1) = 0$ *([8]).*

3.2 Vagueness functions and measure

The idea and construction of a vagueness measure [9] can be derived from the fuzziness measure. Now we will introduce the vagueness measure as a part of the Pliant system.

Definition 8. *The vagueness function* $v : [0, 1] \rightarrow [0, 1]$ *in the Pliant system is*

$$v(x) = \bar{c}(x, \eta(x)) = f^{-1}\left(\frac{1}{2}\left(f(x) + \frac{1}{f(x)}\right)\right) \tag{10}$$

Definition 9. *The normalized vagueness function is*

$$v^*(x) = \frac{1}{v(\nu_0)} v(x), \tag{11}$$

where ν_0 *is the fix point of the negation* η *such that* $\eta(\nu_0) = \nu_0$.

Now we will give a list of properties for the vagueness function:

(P1): $v(x) = 0$ iff $x \in \{0, 1\}$ (Sharpness: no vagueness)

(P2): $v(x)/v(\nu_0) = 1$ iff $x = \nu_0$ (Maximality: maximal vagueness)

(P3): $v(x_1) < v(x_2)$ if   (Monotonicity)

$x_1 < x_2$ and $x_1 \leq \nu_0$ or

$x_1 > x_2$ and $x_1 \geq \nu_0$

(P4): $v(x) = v(\eta(x))$ (Symmetry)

We can get a very simple form of the vagueness function if we use the generator function of the Dombi operator [8], namely $f(x) = \left(\frac{1-x}{x}\right)^\alpha$. First, we get the following formula:

$$v_\alpha(x) = \frac{1}{1 + \left(\frac{1}{2}\left(\frac{1-x}{x}\right)^\alpha + \frac{1}{2}\left(\frac{x}{1-x}\right)^\alpha\right)^{\frac{1}{\alpha}}}. \tag{12}$$

In the Dombi operator case the negation η in the Pliant system is the standard negation $\eta(x) = 1 - x$ (based on Definition 6) and therefore $\nu_0 = \frac{1}{2}$ and $v_\alpha(\nu_0) = \frac{1}{2}$, and we get the following.

Definition 10. *The normalized vagueness function in the Dombi case is*

$$v_\alpha^*(x) = \frac{1}{v_\alpha(\nu_0)}v_\alpha(x) = 2v_\alpha(x) = \frac{2}{1 + \left(\frac{1}{2}\left(\frac{1-x}{x}\right)^\alpha + \frac{1}{2}\left(\frac{x}{1-x}\right)^\alpha\right)^{\frac{1}{\alpha}}}. \tag{13}$$

Let $\alpha = 1$. Then the vagueness function and the normalized vagueness function have the following simple forms:

$$v_1(x) = 2x(1-x) \qquad \text{or} \qquad v_1^*(x) = 4x(1-x) \tag{14}$$

While we got the idea of the vagueness measure from the fuzziness measures, it can be interpreted in a more general way than a measure of fuzziness. Thus, we give a definition which works on a finite set of values from $[0, 1]$.

Definition 11. *Let $\underline{x} = \{x_1, x_2, \dots, x_n\}$ where $x_i \in [0, 1]$, and let v be a vagueness function. Then*

$$\mathcal{V}(\underline{x}) = \frac{1}{n}\sum_{i=1}^{n} v(x_i) \tag{15}$$

is a vagueness measure defined by v.

Definition 12. *Let $\mathcal{V}_\alpha^*$ denote the vagueness measure defined by v_α^*, where v_α^* is the normalized vagueness function in the Dombi case.*

3.3 Relationship between the vagueness measure and the standard deviation

As we showed earlier in [9], the vagueness measure can be interpreted as a fuzziness measure. In this subsection we provide a small example which shows that if the

vagueness measure is applied to a distribution function, it is closely related to the variance of the given population.

In our example, three histograms were generated using three Gaussians with the same mean and different standard deviations (the discrete histogram functions were generated by a simple sampling process). The cumulative distribution functions were also calculated, all functions can be seen in Figure 1. The dashed, the solid and the dotted line histograms were created using 15, 30 and 45 as standard deviation values, respectively. The corresponding vagueness values are 16.92, 33.85 and 49.96 (we used $\mathcal{V}_1^*$ as the vagueness measure), which are related to the former standard deviations. A short explanation is the following. The less the standard deviation, the sharper the distribution function is (see in in Figure 1), and more values are closer to 0 or 1, which leads a smaller vagueness measure value. Table 1 contains more vagueness measure values calculated for the example distribution functions using different α values as measure parameters.

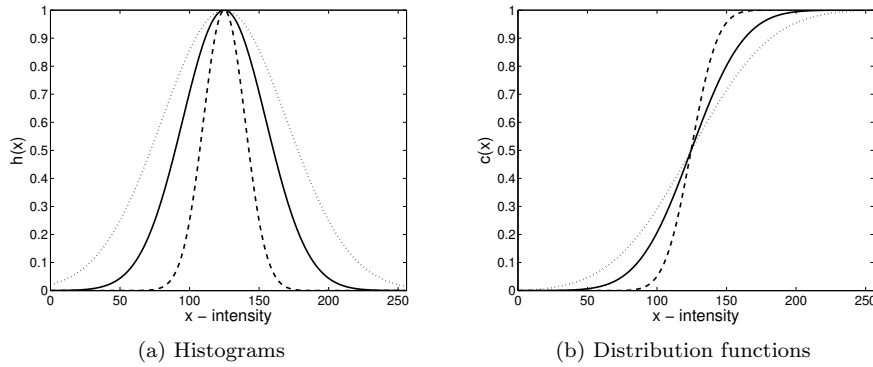(a) Histograms   (b) Distribution functions

Figure 1: Example histograms and cumulative distribution functions. The dashed, the solid and the dotted line histograms were created using Gaussians with the standard deviation values 15, 30 and 45, respectively.

3.4 Relationship between the vagueness and weak ignorance functions

We have to mention the similarity between the vagueness functions and weak ignorance functions which were proposed in [25] and have almost the same properties, although they have different constructions. Weak ignorance functions can be defined based on ignorance functions [4].

Definition 13. *A continuous mapping $G\colon [0,1]^2 \to [0,1]$ is an ignorance function such that:*

σ / α	15.000	30.000	45.000
0.125	32.658	65.218	90.927
0.250	25.777	51.540	74.609
0.500	20.551	41.100	60.360
1.000	16.923	33.846	49.958
2.000	14.658	29.317	43.350
4.000	13.371	26.743	39.570
8.000	12.684	25.369	37.546

Table 1: Relationship between the vagueness measure and the standard deviation. The table contains the vagueness measure values computed for the example distribution functions (with different standard deviation values, appeared in the first row) where the measure has different α values (appeared in the first column).

(G1) $G(x,y) = G(y,x) \ \forall x, y \in [0,1]$
(G2) $G(x,y) = 0 \ iff \ x = 1 \ or \ y = 1$
(G3) $G(0.5, 0.5) = 1$
(G4) G is decreasing in $[0.5,1]^2$
(G5) G is increasing in $[0,0.5]^2$

Theorem 1. *Let* $G \colon [0,1]^2 \to [0,1]$ *be an ignorance function. The continuous function* $g \colon [0,1] \to [0,1]$ *given by*

$$g(x) = G(x, 1-x)$$

is a weak ignorance function which satisfies:

(g1) $g(x) = g(1-x) \ \forall x \in [0,1]$
(g2) $g(x,y) = 0 \ iff \ x \in \{0,1\}$
(g3) $g(0.5) = 1$

The similarities between the vagueness and weak ignorance functions are obvious. The $(g2)$ property is the same as the $(P1)$ property. In the case of the $(g1)$ property, it can be seen that the vagueness function is more general because it expresses the $(P4)$ symmetry using a general negation, while the standard strong negation $N(x) = 1 - x$ appears in the $(g1)$ property. The $(P2)$ property states that a normalized vagueness function gets its maximum $(= 1)$ at the fixpoint of its negation. If we suppose again that in $(g1)$ the strong negation appears with the fixpoint 0.5, then $(g3)$ tells us that at this fixpoint g has its maximum value. It is the same as in $(P2)$ if we restrict it to the strong negation. The monotonicity is not required for weak ignorance functions explicitly.

4 Thresholding based on vagueness measure

Now we present a new global thresholding algorithm which uses the vagueness measure and generalizes the idea of Otsu's method [22]. Although this algorithm works on a grayscale image it can be generalized for the case of real values because it only needs the distribution functions of the populations. First, consider the two vectors of the cumulative distribution function values of the background and object for a given threshold t:

$$\underline{\mathbf{C}}_B^{(t)} = \{c_B^{(t)}(0), c_B^{(t)}(1), \ldots, c_B^{(t)}(t)\}$$
$$\underline{\mathbf{C}}_O^{(t)} = \{c_O^{(t)}(t+1), c_O^{(t)}(t+2), \ldots, c_O^{(t)}(L-1)\}$$

Similarly to Otsu's thresholding criterion [22] which minimizes the intra-class variance, we construct a new method which minimizes the joint vagueness measure of the background and the object in order to yield two well separated population of intensities.

Definition 14. *Let $\underline{\boldsymbol{C}}_B^{(t)}$ and $\underline{\boldsymbol{C}}_O^{(t)}$ the vectors of the cumulative distribution function values of the background and object for a given threshold t. Then the joint vagueness measure belonging to these vectors and the threshold t is defined by the following formula:*

$$p_B^{(t)} \cdot \mathcal{V}_{\alpha_1}^* \left(\underline{\boldsymbol{C}}_B^{(t)} \right) + p_O^{(t)} \cdot \mathcal{V}_{\alpha_2}^* \left(\underline{\boldsymbol{C}}_O^{(t)} \right). \tag{16}$$

With the notations above we can express the new thresholding criterion in the terms of the following optimization problem:

$$t^* = argmin_t \left(p_B^{(t)} \cdot \mathcal{V}_{\alpha_1}^* \left(\underline{\mathbf{C}}_B^{(t)} \right) + p_O^{(t)} \cdot \mathcal{V}_{\alpha_2}^* \left(\underline{\mathbf{C}}_O^{(t)} \right) \right), \tag{17}$$

where t^* is the best threshold corresponding to the minimum joint vagueness measure. The apriori probabilities of the background and the object balance the effect of the vagueness measures if the two population do not have the same size. The pseudo-code of the algorithm can be seen below (Algorithm 1: Pliant Thresholding Algorithm (PTA)). The method requires the histogram h of the image and the parameters of the vagueness measures (α_1, α_2) as the input. The first step can be carried out by using only the histogram. In the second step the calculated values must be used and the vagueness values. The algorithm chooses the threshold t^* which belongs to the lowest joint vagueness measure.

An example image and the running results of the PTA can be seen in Figure 2. Figure 3 contains an example where the PTA was applied on the same image using different values of α_1, α_2.

5 Experimental results

In order to validate the performance of the PTA, it was applied to a set of synthetic images and to a set of standard images (which are widely used in the literature).
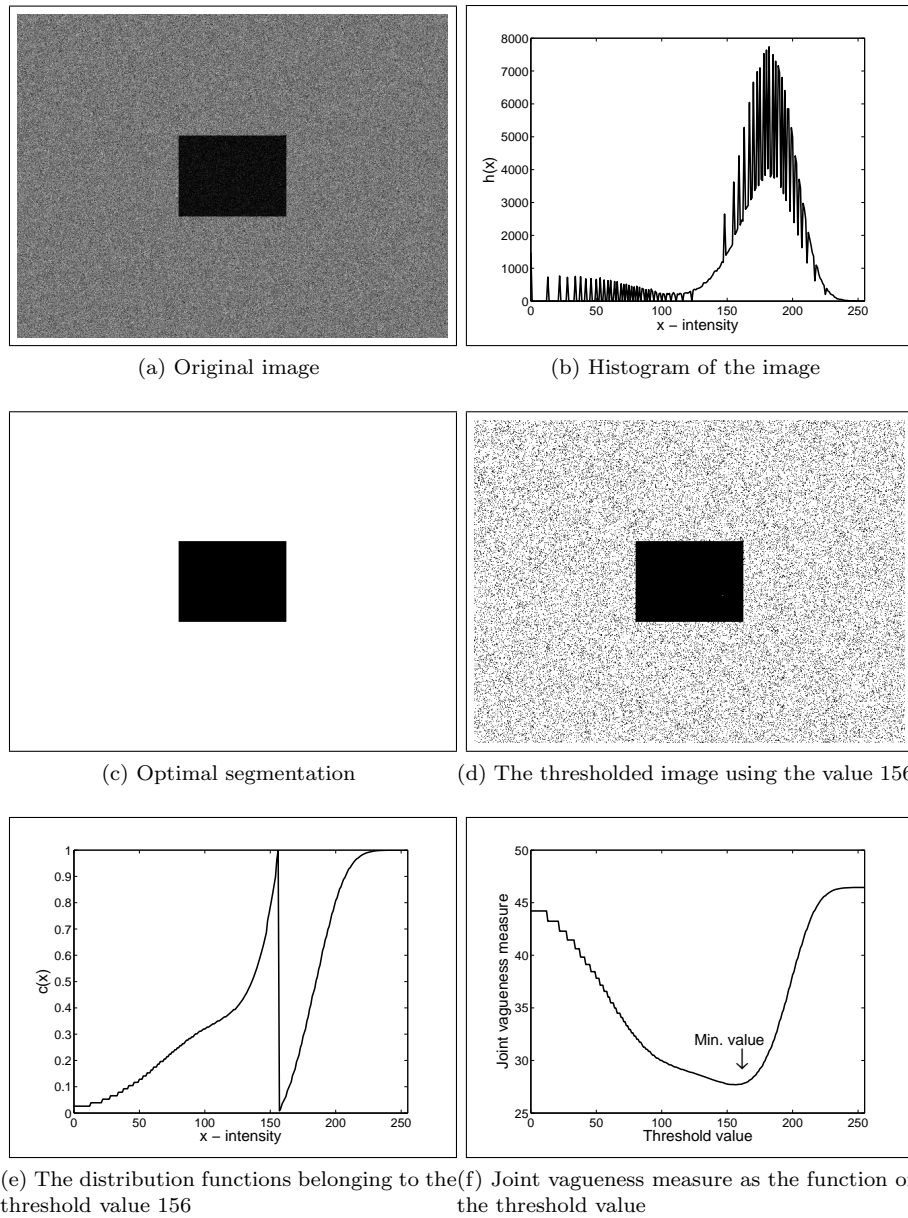
(a) Original image

(b) Histogram of the image

(c) Optimal segmentation

(d) The thresholded image using the value 156

(e) The distribution functions belonging to the threshold value 156

(f) Joint vagueness measure as the function of the threshold value

Figure 2: An example image and the results of the PTA. The algorithm chose the threshold value 156.

Algorithm 1 Pliant Thresholding Algorithm (PTA)

Require: histogram h, parameters of the two vagueness measures α_1, α_2

(a) compute $p_B^{(t)}, p_O^{(t)}, c_B^{(t)}, c_O^{(t)}$ for all $t \in [0, L-1]$

(b) compute joint vagueness measures for all $t \in [0, L-1]$

 (i) compute value $\mathcal{V}_{\alpha_1}^* \left(\underline{\mathbf{C}}_B^{(t)} \right)$

 (ii) compute value $\mathcal{V}_{\alpha_2}^* \left(\underline{\mathbf{C}}_O^{(t)} \right)$

 (iii) compute the joint vagueness measure

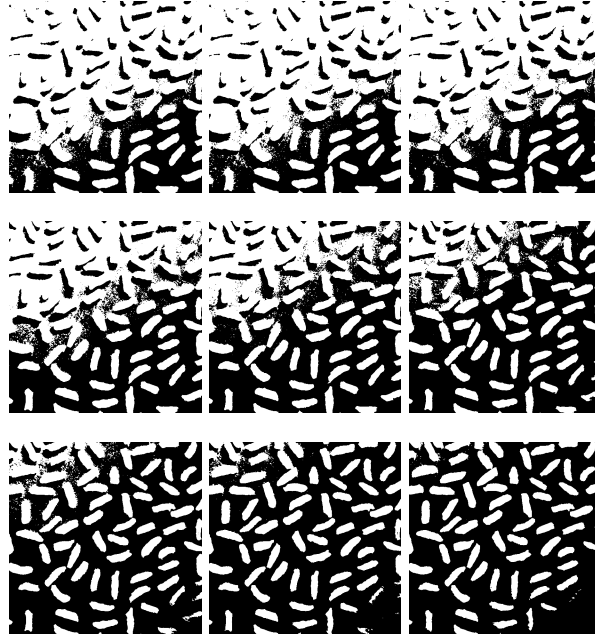(c) take t^* as the best threshold corresponding to the minimum joint vagueness measure

---

Figure 3: Applying of the PTA using different sets of (α_1, α_2): $(0.0625, 16.0)$, $(0.125, 8.00)$, $(0.25, 4.00)$, $(0.50, 2.00)$, $(1.00, 1.00)$, $(2.00, 0.50)$, $(4.00, 0.25)$, $(8.00, 0.125)$, $(16.00, 0.0625)$, from left to right, top to bottom.

We compared the PTA (using the parameters $\alpha_1 = \alpha_2 = \frac{1}{2}$) with some other well-known and commonly used thresholding methods, namely Otsu's method [22], Huang-Wang algorithm [13], Kittler's method [16], Li's method [20], and Area algorithm [3] (with $\phi_1(x) = \phi_2(x) = x$). There are many approaches available for comparing thresholding algorithms in the literature [19, 24, 26, 33]. Here we have applied the misclassification error [30] as the performance criterion which is the following.

Definition 15. *Let B_O and F_O denote the known background and object pixel sets belonging to the optimal segmentation, and let B_T and F_T be the background and object area pixels in the result image thresholding by t, respectively. Then the misclassification error which belongs to a given threshold t is computed by the following formula:*

$$ME = 1 - \frac{|B_O \cap B_T| + |F_O \cap F_T|}{|B_O| + |F_O|}. \tag{18}$$

5.1   Comparison on the sets of synthetic images

Now, we introduce the creating process of the synthetic images which were used in the first evaluation. The first (basic) image set was created as follows. An image consists of a square on the center of the image which is the object and the other surrounding pixels compose the background (we do not need more complicated shapes because the algorithm works on the image's histogram). The object and the background were generated by two Gaussians (with different means but with not necessarily different standard deviations). The means were picked up from the set $\{0, 30, 50, 120, 150, 200\}$ while for the deviations we used the set $\{10, 20, 30\}$ (not all of the possible selections were generated). So we got a synthetic set with 250 images. We applied different distortions on the synthetic images in order to yield more test cases: we used additive Gaussian noise (with different standard deviations) and impulsive salt-and-pepper noise (with different amount of corrupted pixels). Each distortion with 4 different parameters were used on the original images so the total number of the images was $250 + 250 \cdot 2 \cdot 4 = 2250$. Some examples can be seen in Figure 4.
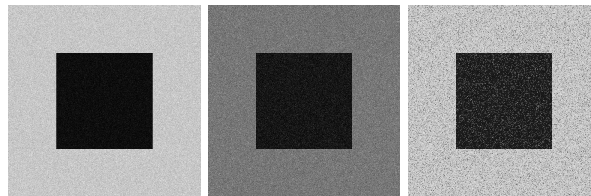

Figure 4: Synthetic image examples (from the left to the right): standard image without any distorsions, image with Gaussian noise, and image contaminated by salt-and-pepper noise.

The summarized results can be found in Table 2. The rows belong to different synthetic sets: the first contains the misclassification errors on the original images while the next four show the results on the images with additive Gaussian noise, and the last four give the values for the salt-and-pepper noise contaminated images. As it can be seen the PTA and the Area algorithm have almost the same error values (small differences occur in the the cases of the salt-and-pepper noise contaminated images) and both perform better than the other studied methods.

Type	Otsu	Huang-W.	Kittler	Li	Area	PTA ($\alpha = 0.5$)
B	0.049	0.049	0.071	0.053	0.045	0.045
G2	0.049	0.049	0.080	0.054	0.045	0.045
G5	0.051	0.055	0.084	0.056	0.048	0.048
G10	0.058	0.068	0.113	0.065	0.059	0.059
G15	0.070	0.093	0.164	0.079	0.076	0.076
SP1	0.054	0.052	0.077	0.058	0.049	0.048
SP3	0.071	0.061	0.078	0.072	0.058	0.056
SP5	0.099	0.072	0.089	0.090	0.069	0.065
SP10	0.124	0.098	0.109	0.136	0.096	0.108

Table 2: Misclassification errors on synthetic images. Every row belongs to a set of synthetic images: B means the original synthetic images, G denotes the images with additive Gaussian noise where the number is the standard deviation σ, and SP refers to the salt-pepper noise contaminated images where the number gives the percentage of the modified pixels.

5.2   Comparison on the set of standard images

The second evaluation of the algorithm was investigated by using 20 classical and bimodal test images which are taken from a collection of Carnegie Mellon University (http://www.cs.cmu.edu/~cil/v-images.html). The images can be seen in Figure 5 while the ground-truth images appear in Figure 6 where the ground-truth thresholds were set manually.

The missclassification error values are presented in Table 3 below. These values are the average measures calculated for the 20 test images. In this case, the PTA gives a slightly worse result than Otsu, but Li's algorithm and the Area perform very similarly. We should mention here that the measure ME suffers from the subjectivity of the human expert or observer who sets the ground-truth thresholds, so these differences could be regardarded as insignificant. However we have chosen the ground-truth threshold instead of creating a gold standard mask because we compare global thresholding algorithms, and we think in this case that the ground-truth threshold value gives the technical maximum for these algorithms at least from ME aspect.
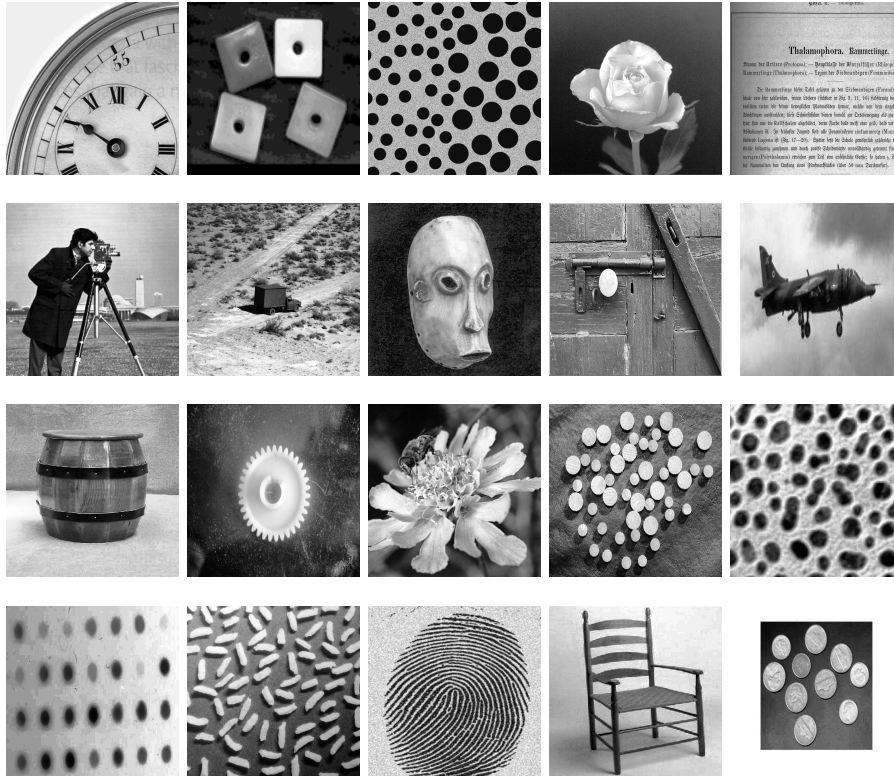
Figure 5: Original images

Images	Otsu	Huang-W.	Kittler	Li	Area	PTA ($\alpha = 0.5$)
Standard	0.085	0.135	0.185	0.101	0.106	0.095

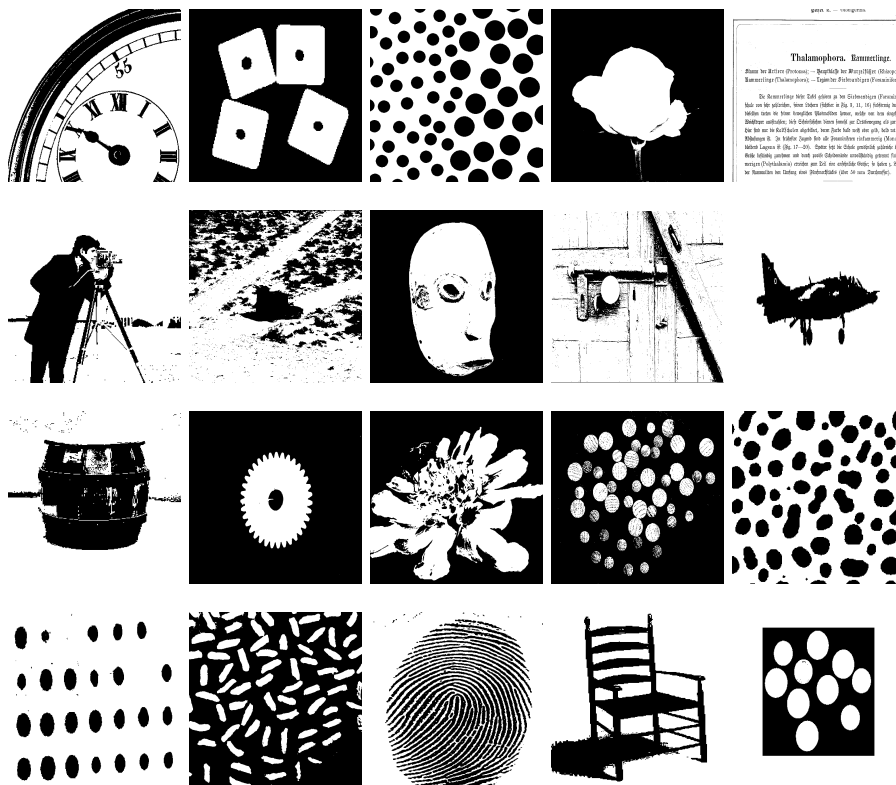Table 3: Misclassification errors on standard images.

Figure 6: Ground-truth images

6   Conclusions

In this study we provided a new interpretation of the vagueness measure which assigns a value to a distribution function that is similar to the population's standard deviation. Then we gave an application using this property (in the form of a global thresholding algorithm). The proposed method performs as well as or better than the classical and state-of-the-arts methods. One advantage of this approach is that the algorithm can be parameterized, hence it can be adapted to a given problem. In the future, we would like to try other conjunction operators taken from the Pliant system (e.g. a weighted conjunction) that have more parameters. It is an interesting question of how the parameters should be optimized such that it fits the given problem. Anyway, it would be useful to know what classes of problems the vagueness measure can be applied to.

7   Acknowledgments

References

[1] Alsina, C., Schweizer, B., and Frank, M.J. *Associative functions: triangular norms and copulas.* Word Scientific Publishing, 2006.

[2] Beliakov, G., Pradera, A., and Calvo, T. *Aggregation Functions: A Guide for Practitioners.* Studies in Fuzziness and Soft Computing. Vol. 221., Springer, 2007.

[3] Bustince, H., Barrenechea, E., and Pagola, M. Image thresholding using restricted equivalence functions and maximizing the measures of similarity. *Fuzzy Sets and Systems*, 158(5):496–516, 2007.

[4] Bustince, H., Pagola, M., Barrenechea, E., Fernandez, J., Melo-Pinto, P., Couto, P., Tizhoosh, HR, and Montero, J. Ignorance functions. An application to the calculation of the threshold in prostate ultrasound images. *Fuzzy Sets and Systems*, 161(1):20–36, 2010.

[5] Calvo, T., Mayor, G., and Mesiar, R., editors. *Aggregation Operators.*, volume 97 of *New Trends and Applications Studies in Fuzziness and Soft Computing.* Physica-Verlag, Heidelberg, 2002.

[6] Chaira, T. and Ray, AK. Segmentation using fuzzy divergence. *Pattern Recognition Letters*, 24(12):1837–1844, 2003.

[7] Chaira, T. and Ray, AK. Threshold selection using fuzzy set theory. *Pattern Recognition Letters*, 25(8):865–874, 2004.

[8] Dombi, J. Demorgan systems with an infinitely many negations in the strict monotone operator case. *Information Sciences*, 181:1440–1453, 2011.

[9] Dombi, J. Fuzziness measure in the pliant system: The vagueness measure. *Acta Technica Jaurinensis*, 4(1), 2011.

[10] Dombi, J. On certain class of aggregation operator. *Information Sciences, Under review process*, 2011.

[11] Dombi, J. Pliant operator system. *Recent Advances in Intelligent Engineering Systems, Studies in Computational Intelligence, Springer*, 378:31–58, 2012, under print.

[12] Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E. *Aggregation Functions.* Encyclopedia of Mathematics and Its Applications 127, Cambridge University Press, Cambridge, 2009.

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

[14] Jawahar, CV, Biswas, PK, and Ray, AK. Investigations on fuzzy thresholding based on fuzzy clustering. *Pattern Recognition*, 30(10):1605–1613, 1997.

[15] Kapur, JN, Sahoo, PK, and Wong, AKC. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing*, 29(3):273–285, 1985.

[16] Kittler, J. and Illingworth, J. Minimum error thresholding. *Pattern Recognition*, 19(1):41–47, 1986.

[17] Klement, E.P., Mesiar, R., and Pap, E. *Triangular norms.* Dordrecht:Kluwer, 2000.

[18] Lee, S.U., Yoon Chung, S., and Park, R.H. A comparative performance study of several global thresholding techniques for segmentation. *Computer Vision, Graphics, and Image Processing*, 52(2):171–190, 1990.

[19] Levine, M.D. and Nazif, A.M. Dynamic measurement of computer generated image segmentations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2):155–164, 2009.

[20] Li, CH and Tam, PKS. An iterative algorithm for minimum cross entropy thresholding. *Pattern Recognition Letters*, 19(8):771–776, 1998.

[21] Mesiar, R., Kolesárová, A., Calvo, T., and Komornková, M. *A Review of Aggregation Functions.* Fuzzy Sets and Their Extension: Representation, Aggregation and Models. Studies in Fuzziness and Soft Computing, Vol. 220, 2008.

[22] Otsu, N. A threshold selection method from gray-level histograms. *Automatica*, 11:285–296, 1975.

[23] Ridler, TW and Calvard, S. Picture thresholding using an iterative selection method. *Systems, Man and Cybernetics, IEEE Transactions on*, 8(8):630–632, 2007.

[24] Sahoo, PK, Soltani, S., and Wong, AKC. A survey of thresholding techniques. *Computer Vision, Graphics, and Image Processing*, 41(2):233–260, 1988.

[25] Sanz, Josean, Fernndez, Alberto, Sola, Humberto Bustince, and Herrera, Francisco. A genetic tuning to improve the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets: Degree of ignorance and lateral position. *Int. J. Approx. Reasoning*, pages 751–766, 2011.

[26] Sezgin, M. and Sankur, B. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1):146–168, 2004.

[27] Tizhoosh, H.R. Image thresholding using type II fuzzy sets. *Pattern Recognition*, 38(12):2363–2372, 2005.

[28] Wang, Q., Chi, Z., and Zhao, R. Image thresholding by maximizing the index of nonfuzziness of the 2-D grayscale histogram. *Computer Vision and Image Understanding*, 85(2):100–116, 2002.

[29] Weszka, J.S. A survey of threshold selection techniques. *Computer Graphics and Image Processing*, 7(2):259–265, 1978.

[30] Yasnoff, W.A., Mui, J.K., and Bacus, J.W. Error measures for scene segmentation. *Pattern Recognition*, 9(4):217–231, 1977.

[31] Zadeh, L.A. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.

[32] Zadeh, L.A. *Fuzzy sets and their applications to cognitive and decision processes.* Academic Press, New York, 1975.

[33] Zhang, Y.J. A survey on evaluation methods for image segmentation. *Pattern Recognition*, 29(8):1335–1346, 1996.