A New Measure of Fuzzy Directed Divergence and Its Application in Image Segmentation

P.K Bhatia

Department of Mathematics, DCR University of Science and Technology, Murthal-131039 (Haryana), India
Email: bhatiapk@rediffmail.com

Surender Singh

School of Mathematics, Shri Mata Vaishno Devi University, Sub post office, Katra-182320 (J & K), India
Email: surender1976@gmail.com

Abstract— An approach to develop new measures of fuzzy directed divergence is proposed here. A new measure of fuzzy directed divergence is proposed, and some mathematical properties of this measure are proved. The application of fuzzy directed divergence in image segmentation is explained. The proposed technique minimizes the fuzzy divergence or the separation between the actual and ideal thresholded image.

Index Terms— Aggregation, Divergence, Gamma Distribution, Thresholding

I. Introduction

In 1948 Shannon [1] advanced information theory (IT) as a new branch of mathematics and a powerful tool for understanding the intricacies of the communication process. Because of the limitations of Shannon measure in certain situations Renyi [2] took the first step and generalized the Shannon measure. After Renyi many generalized measures were developed for different situations. Kullback and Leibler [3] introduced the measure of discrimination between two probability distributions, one is ideal another is observed. Literature on the development of divergence measures has expanded considerably in the last two decades of 20th century. Taneja [4], Besseville [5], Esteban and Morales [6] reviewed the development of generalized information and divergence measures.

The concept of fuzziness introduced by Zadeh [7] revolutionized research and development in the area. De-Luca and Termini [8] defined the measure of fuzzy entropy corresponding to Shannon [1] measure of entropy. Bhandari and Pal [9] defined measures of fuzzy entropy corresponding to Renyi [2] entropy and measure of fuzzy directed divergence corresponding to Kullback Leibler [3] divergence measure. Literature on the development of divergence measures has expanded considerably in recent years. Bhatia and Singh [10]

presented a survey of fuzzy information and divergence measures.

Thresholding, a popular tool for image segmentation, is used here for extracting the objects from a picture. If the objects are clearly distinguishable from the background, the threshold values for segmentation can be chosen at the valley points of the multimodal histogram.

Otsu [11] selected the threshold so as to maximize the class separatability, which was based on within class variance, between class variance and total variance of gray levels. A number of excellent investigations on various thresholding techniques are reported in the literature. Kapur et al. [12], Sahoo et al. [13,14], Sahoo and Wong [15], and Brink and Pendcock [16] used information theoretic measures to threshold an image.

Fuzzy set theory is applied to image thresholding to partition the image into meaningful regions. The application of fuzzy notions in image processing has gained importance because of several reasons like (i) imprecision of the gray levels of an image, (ii) ambiguity in some definitions, such as boundaries between the regions, or region textures etc. Thus it is possible to allow the segments to be several fuzzy subsets of the image. The measures like entropy, index of fuzziness and index of non fuzziness can be used as objective functions which may be optimized for image segmentation. Index of fuzziness represents the average amount of fuzziness in an image by measuring the distance between the fuzzy property of an image and its nearest two-tone version. The index of non-fuzziness indicates the amount of non-fuzziness in an image by taking an absolute difference between the fuzzy property of an image and its complement.

Several researchers used fuzzy based thresholding techniques. Pal and Dasgupta [17] introduced a concept of spectral fuzzy sets for finding the membership value and then segmented the image. Huang and Wang [18] assigned the membership value by taking the reciprocal of the absolute difference of pixel and the mean of the region to which that pixel

belongs. Ramar et al. [19] used the neural network for selecting the best threshold using various fuzzy measures viz., linear and quadratic indices of fuzziness, logarithmic and exponential entropy. Cheng and Chen [20] used fuzzy homogeneity vectors and fuzzy co-occurrence matrix for image thresholding.

In this paper a new approach to develop probabilistic divergence measures and thereby measures of fuzzy directed divergence is presented. A new measure of fuzzy directed divergence is introduced and some of its properties studied. A new method based on minimization of fuzzy directed divergence using Gamma distribution to determine the membership function of pixels of an image introduced by Chaira and Ray [21] is applied in context with the newly developed measure of fuzzy directed divergence. The proposed methodology involves the minimization of the divergence between the pixels in an ideally thresholded image and actually thresholded image.

This paper is organized as follows. Section II addresses the preliminaries. Section III presents an approach to develop probabilistic divergence measures . In section IV, a new measure of fuzzy directed divergence is presented. Section V deals the application of new measure of fuzzy directed divergence in image segmentation, and section VI contains concluding remarks.

II. Preliminaries

2.1 Information Measure

The measure of information was defined Claude E. Shannon in his treatise paper [1] in 1948.

$$H(P) = \sum_{i=1}^{n} p_i \log p_i , \qquad P \in \Gamma_n \tag{1}$$

Where

$$\Gamma_n = \{P = (p_1, p_2, ... p_n) / p_i \geq 0, \sum_{i=1}^{n} p_i = 1 ; n \geq 2\}$$

is the set of all complete finite discrete probability distributions. To improve upon the weakness of Shannon's measure in certain situations Renyi [2] took the first step and proposed a parametric measure of information

$$H_\alpha(P) = \frac{1}{1-\alpha} \log (\sum_{i=1}^{n} p_i^\alpha), \alpha \neq 1, \alpha > 0 \tag{2}$$

2.2 Divergence Measure

The relative entropy or directed divergence is a measure of the distance between two probability distributions. In statistics, it arises as the expected logarithm of the likelihood ratio. The relative entropy

D(P,Q) is the measure of inefficiency of assuming that the distribution is Q when the true distribution is P . For example, if we knew the true distribution of the random variable, then we could construct a code with average description length H(P). If, instead, we used the code for a distribution Q, we would need H(P) + D(P,Q) bits on the average to describe the random variable. The relative entropy or Kullback-Leibler distance Kullback and Leibler [3] between two probability distributions is defined as

$$D(P,Q) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} \tag{3}$$

A correct measure of directed divergence must satisfy the following postulates:

a. $D(P,Q) \geq 0$

b. $D(P,Q) = 0$ iff $P = Q$

c. $D(P, Q)$ is a convex function of both

$$P = (p_1, p_2, ..., p_n) \text{ and } Q = (q_1, q_2, ..., q_n)$$

If in addition, symmetry and triangle inequality is also satisfied by D(P,Q) then it is called a **distance measure**. Properties (a)-(c) are essential to define a new measure of directed divergence. A parametric measure of directed divergence can also be characterized in terms of its parameter(s).

2.3 Fuzzy Sets

Definition [7]. Let a universe of discourse X = {x_1, x_2, x_3 ... x_n} then a fuzzy subset of universe X is defined as

A = {(x; μ_A(x)) / x $\in$ X; μ_A(x): X $\rightarrow$ [0; 1]}

Where μ_A(x): X $\rightarrow$ [0; 1] is a membership function defined as follow

$$\mu_A(x) = \begin{cases} 0 & \text{if x does not belong to A and there is no ambiguity} \\ 1 & \text{if x belong to A and there is no ambiguity} \\ 0.5 & \text{if there is maximum ambiguity whether x belongs to A or not} \end{cases}$$

In fact μ_A(x) associates with each x $\in$ X a grade of membership of the set A. Some notions related to fuzzy sets [7].

Containment; A $\subseteq$ B $\Leftrightarrow$ μ_A(x) $\leq$ μ_B(x) for all x $\in$ X

Equality; A = B $\Leftrightarrow$ μ_A(x) = μ_B(x) for all x $\in$ X

Compliment; $\overline{A}$ = Compliment of A $\Leftrightarrow$

$\mu_{\bar{A}}(x) = 1-\mu_A(x)$ for all $x \in X$

Union; $A \cup B$ = Union of A and B $\Leftrightarrow$

$\mu_{A \cup B}(x) = \max.\{\mu_A(x), \mu_B(x)\}$ for all $x \in X$

Intersection; $A \cap B$ = Intersection of A and B $\Leftrightarrow \mu_{A \cap B}(x) = \min.\{\mu_A(x), \mu_B(x)\}$ for all $x \in X$

Product; AB = Product of A and B $\Leftrightarrow$

$\mu_{AB}(x) = \mu_A(x)\mu_B(x)$ for all $x \in X$

Sum; $A \oplus B$ = Sum of A and B $\Leftrightarrow$

$\mu_{A \oplus B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)$ for all $x \in X$

2.4 Fuzzy directed divergence

Definition[22]**.** Let a universal set X and F (X) be the set of all fuzzy subsets .A mapping D:F (X) × F (X) $\rightarrow$ R is called a **divergence** between fuzzy subsets if and only if the following axioms hold:

a. D (A, B)

b. D (A, B) =0 if A=B

c. $\max.\{D(A \cup C, B \cup C), D(A \cap C, B \cap C)\} \le D(A,B)$ for

any A, B, C $\in$ F(X)

Instead of axiom (c) if D (A, B) is convex in A and B even then it is a valid measure of divergence.

Bhandari and Pal [9] defined measure of fuzzy directed divergence corresponding to (3) as follow:

$$D(A,B) = \sum_{i=1}^{n} \mu_A(x_i) \log \frac{\mu_A(x_i)}{\mu_B(x_i)} + \sum_{i=1}^{n} (1-\mu_A(x_i)) \log \frac{(1-\mu_A(x_i))}{(1-\mu_B(x_i))}$$
(3A)

Definition [23]. Let a universal set X and F (X) be the set of all fuzzy subsets .A mapping D: F (X) ×F (X) $\rightarrow$R is called a distance measure on F(X) if and only if the following axioms hold:

a. D (A, B) =D (B, A)

b. D (A, B) =0 if A=B

c. $D(T,\bar{T}) =_{Max\ A,B \in F(X)} D(A,B) \quad \forall T \in P(X)$

d. $\forall A, B, C \in F(X)$ if $A \subset B \subset C$, then

$D(A,B) \le D(A,C)$ and $D(B,C) \le D(A,C)$

2.5 Aggregation operations

The aggregation operation on fuzzy sets is the operations by which several fuzzy sets are combined to produce a single set. e.g fuzzy union and fuzzy intersection are special cases of aggregation operations.

Defnition [10]. An aggregation operation is defined by the function $M:[0,1]^n \rightarrow [0,1]$ verifying

1. M(0,0,0,...0) = 0 , M(1, 1, 1,...,1) = 1 (Boundary Conditions)

2. M is Monotonic in each argument. (Monotonicity)

The use of monotone functions is justified in many decision making contexts, since it ensures consistency and reliability. The boundary conditions here are specified with the assumption that inputs are provided on the unit interval, however in certain cases, inputs naturally expressed on different intervals can be scaled appropriately. If n=2 then M is called a binary aggregation operation.

Aggregation functions are classed depending on their behavior relative to the inputs. The most commonly used in application are averaging functions, which are usually interpreted as being representative of a given set of inputs or input vector.

III. An Approach to Develop Probabilistic Divergence Measure

Let $U(a,b)$ and $V(a,b)$ be two binary aggregation operators then

$$D(P,Q) = \sum_i |U(p_i,q_i) - V(p_i,q_i)|$$
(4)

Where $P,Q \in \Gamma_n$ as in section II.

is divergence measure.

We have $A^*:[0,1]^2 \rightarrow [0,1]$ such that

$$A^*(a,b) = \frac{a+b}{2}$$
(5)

and $H^*:[0,1]^2 \rightarrow [0,1]$ such that

$$H^*(a,b) = \frac{a^2+b^2}{a+b}$$
(6)

are aggregation operations.

Now using these aggregation operators following divergence measure is defined by Bhatia and Singh [24].

$$D_{H^*A^*}(P,Q) = \sum_{i=1}^{n}\left[\frac{p_i^2 + q_i^2}{p_i + q_i} - \frac{p_i + q_i}{2}\right]$$

$$= \sum_{i=1}^{n}\frac{(p_i - q_i)^2}{2(p_i + q_i)} \tag{7}$$

Taneja [25] has defined several probabilistic divergence measures using this approach. Researchers have defined measures of fuzzy directed divergence corresponding to classical probabilistic divergence measures, a brief review is presented by Bhatia and Singh [24]. In section IV a new measure of fuzzy directed divergence corresponding to (7) is defined.

IV. New Measure of Fuzzy Directed Divergence

The measure of fuzzy directed divergence between two fuzzy sets corresponding to (7) is defined as follows:

$$M_{H^*A^*}^{F}(A,B) =$$
$$\sum_{i=1}^{n}\frac{(\mu_A(x_i) - \mu_B(x_i))^2}{2}\left[\frac{1}{\mu_A(x_i) + \mu_B(x_i)} + \frac{1}{2 - \mu_A(x_i) - \mu_B(x_i)}\right] \tag{8}$$

Theorem 1. $M_{H^*A^*}^{F}(A,B)$ s a valid measure of fuzzy directed divergence.

Proof. In order to prove that $M_{H^*A^*}^{F}(A,B)$ s a valid measure of divergence three axioms (a), (b) and (c) of fuzzy directed divergence must be satisfied.

a. From definition of $M_{H^*A^*}^{F}(A,B)$ it is obvious that

$$M_{H^*A^*}^{F}(A,B) = M_{H^*A^*}^{F}(B,A)$$

and

$$M_{H^*A^*}^{F}(A,A) = 0$$

b. To prove axiom (c) we divide the X, the universe of discourse into six subsets as follow:

$$W_1 = \{x \mid x \in X, \mu_A(x) \le \mu_B(x) \le \mu_C(x)\}$$

$$W_2 = \{x \mid x \in X, \mu_A(x) \le \mu_C(x) < \mu_B(x)\}$$

$$W_3 = \{x \mid x \in X, \mu_B(x) < \mu_A(x) \le \mu_C(x)\}$$

$$W_4 = \{x \mid x \in X, \mu_B(x) \le \mu_C(x) < \mu_C(x)\}$$

$$W_5 = \{x \mid x \in X, \mu_C(x) < \mu_A(x) \le \mu_B(x)\}$$

$$W_6 = \{x \mid x \in X, \mu_C(x) < \mu_B(x) < \mu_A(x)\}$$

In W_1,

$A \cup C =$ Union of A and C

$$\Leftrightarrow \mu_{A \cup C}(x) = \max.\{\mu_A(x), \mu_C(x)\} = \mu_C(x)$$

$B \cup C =$ Union of B and C

$$\Leftrightarrow \mu_{B \cup C}(x) = \max.\{\mu_B(x), \mu_C(x)\} = \mu_C(x)$$

$A \cap C =$ Intersection of A and C

$$\Leftrightarrow \mu_{A \cap C}(x) = \min.\{\mu_A(x), \mu_C(x)\} = \mu_A(x)$$

$B \cap C =$ Intersection of B and C

$$\Leftrightarrow \mu_{B \cap C}(x) = \min.\{\mu_B(x), \mu_C(x)\} = \mu_B(x)$$

Therefore from Eq.(8), we have

$$M_{H^*A^*}^{F}(A \cup C, B \cup C) = 0$$

and

$$M_{H^*A^*}^{F}(A \cap C, B \cap C) = M_{H^*A^*}^{F}(A,B)$$

Hence in set, W_1

$$\max.\{M_{H^*A^*}^{F}(A \cup C, B \cup C), M_{H^*A^*}^{F}(A \cap C, B \cap C)\} \le M_{H^*A^*}^{F}(A,B)$$

By similar calculations it can be observed that above inequality also holds on sets W_2, W_3, W_4, W_5 and W_6.

Thus $M_{H^*A^*}^{F}(A,B)$ is a valid measure of fuzzy directed divergence.

Theorem 2. Following properties can be verified for $M_{H^*A^*}^{F}(A,B)$.

1. $M_{H^*A^*}^{F}(A \cup B, A \cap B) = M_{H^*A^*}^{F}(A,B)$

2. $M_{H^*A^*}^{F}(A, \overline{A}) = n$ when A is a crisp set i.e when $\mu_A(x) = 0$ or 1.

3. $M_{H^*A^*}^{F}(A,B) = M_{H^*A^*}^{F}(\overline{A}, \overline{B}) \quad \forall A, B \in F(X)$

Proof. we divide the X, the universe of discourse into two subsets as follows:

$$W_1 = \{x \mid x \in X, \mu_A(x) \le \mu_B(x)\}$$

$$W_2 = \{x \mid x \in X, \mu_A(x) > \mu_B(x)\}$$

The proof can be obtained by performing similar calculations as in theorem 1 on sets W_1 and W_2.

Theorem3. $M_{H^*A^*}^F$ is not a distance measure on F(X).

Proof. From definition of $M_{H^*A^*}^F(A,B)$ it is obvious that

$$M_{H^*A^*}^F(A,B) = M_{H^*A^*}^F(B,A)$$

and

$$M_{H^*A^*}^F(A,A) = 0$$

Next we show that,

$$M_{H^*A^*}^F(T,\overline{T}) = {}_{Max\,A,B \in F(X)} M_{H^*A^*}^F(A,B) \quad \forall T \in P(X)$$

Differentiating Eq.(8) partially with respect to $\mu_A(x)$ and $\mu_B(x)$ respectively we have

$$\frac{\partial M_{H^*A^*}^F}{\partial \mu_A} = 2\mu_B^2 - \mu_A - 3\mu_B + \mu_A\mu_B$$

$$\frac{\partial M_{H^*A^*}^F}{\partial \mu_B} = 2\mu_A^2 - \mu_B - 3\mu_A + \mu_A\mu_B$$

For points of maxima and minima,

$$\frac{\partial M_{H^*A^*}^F}{\partial \mu_A} = 0 \quad \text{and} \quad \frac{\partial M_{H^*A^*}^F}{\partial \mu_B} = 0$$

We observe that

$M_{H^*A^*}^F$ is minimum when $\mu_A(x) = \mu_B(x)$

and maximum when $\mu_B(x) = 1 - \mu_A(x) = \mu_{\overline{A}}(x)$

Therefore

$$M_{H^*A^*}^F(T,\overline{T}) = {}_{Max\,A,B \in F(X)} M_{H^*A^*}^F(A,B) \quad \forall T \in P(X)$$

Finally we observe that $\forall A,B,C \in F(X)$ if $A \subset B \subset C$, then $M_{H^*A^*}^F(A,B) \leq M_{H^*A^*}^F(A,C)$ and

$$M_{H^*A^*}^F(B,C) \leq M_{H^*A^*}^F(A,C)$$

is not satisfied.

V. Application of new measure of fuzzy directed divergence in image segmentation

5.1 Mathematical modeling of image

An image can be described by a two-dimensional function f(x,y), where (x, y) denotes the spatial coordinate and f(x,y) the feature value at (x, y). Depending on the type of image, the feature value could be light intensity, depth, intensity of radio wave or temperature. A digital image, on the other hand, is a two-dimensional discrete function f(x,y) which has been digitized both in spatial coordinates and magnitude of feature value. We shall view a digital image as a two-dimensional matrix whose row and column indices identify a point, called a pixel, in the image and the corresponding matrix element value identifies the feature intensity level. Here a digital image will be represented as

$$F_{M \times N} = \left[f(x,y)\right]_{M \times N}$$

Where $M \times N$ is the size of the image and f (x. y) $\in G_L = \{0, 1,..., L-1\}$, the set of discrete levels of the feature value. Since the majority of the techniques we are going to discuss in this paper are developed primarily for ordinary intensity images, in our subsequent discussion, we shall usually refer to f(x, y) as gray value (although it could be depth or temperature or intensity of radio wave).

5.2 Gray level thresholding

Thresholding is one of the old, simple and popular techniques for image segmentation. Thresholding can be done based on global information (e.g. gray level histogram of the entire image) or it can be done using local information (e.g. co-occurrence matrix) of the image. Taxt et al.[26] refer to the local and global information based techniques respectively as contextual and non contextual methods. Under each of these schemes (contextual/non-contextual) if only one threshold is used for the entire image then it is called global thresholding. On the other hand, when the image is partitioned into several sub regions and a threshold is determined for each of the sub regions, it is referred to as local thresholding. Thresholding techniques can also be classified as bi-level thresholding and multi-thresholding. In bi-level thresholding the image is partitioned into two regions—object (black) and background (white). When the image is composed of several objects with different surface characteristics (for a light intensity image, objects with different coefficient of reflection, for a range image there can be objects with different depths and so on) one needs several thresholds for segmentation. This is known as multi-thresholding. In such a situation we try to get a set of thresholds (t₁, t₂, ..., tₖ) such that all pixels with ,i = 0, 1, ..., k: constitute the ith region type (t₀ and tₖ₊₁, are taken as 0 and L-1, respectively). Note that thresholding can also be viewed as a classification problem. For example, bi-level segmentation is equivalent to classifying the pixels into two classes: object and background.

If the image is composed of regions with different gray level ranges, i.e. the regions are distinct, the histogram of the image usually shows different peaks, each corresponding to one region and adjacent peaks

are likely to be separated by a valley. For example, if the image has a distinct object on a background, the gray level histogram is likely to be bimodal with a deep valley. In this case, the bottom of the valley (T) is taken as the threshold for object background separation. Therefore, when the histogram has a (or a set of) deep valley(s), selection of threshold(s) becomes easy because it becomes a problem of detecting valleys. However, normally the situation is not like this and threshold selection is not a trivial job. There are various methods available for this. Here we describe thresholding by fuzzy divergence.

5.3 Fuzziness and membership function

Let $X = \{f_{ij}, \mu(f_{ij})\} \quad \forall \ f_{ij} \in X,$ be an image of size $M \times M$ having L levels and f_{ij} be grey level of $(i, j)th$ pixel in X. Let $\mu(f_{ij})$ denote the membership value of $(i, j)th$ pixel in X, where $0 \leq \mu(f_{ij}) \leq 1$ with $\mu(f_{ij}) = 1$ denoting full membership and $\mu(f_{ij}) = 0$ denoting non-membership.

For a given threshold value, the membership function derived from the gamma distribution is proportional to the exponential function of negative of the absolute difference between the mean of the region to which the pixel belongs and the pixel gray level. It is thus obvious that this is inversely proportional to the membership value. Let *count(f)* denote the number of occurrences of the gray level f in the image. Given a certain threshold value t, which separates the object and the background, the average gray level of the background region is given by the relation:

$$\mu_0 = \frac{\sum\limits_{f=0}^{t} f.count(f)}{\sum\limits_{f=0}^{t} count(f)}$$

(9)

and the average gray level of object region is given by

$$\mu_1 = \frac{\sum\limits_{f=t+1}^{L-1} f.count(f)}{\sum\limits_{f=t+1}^{L-1} count(f)}$$

(10)

The membership function of each pixel in the image depends on its affinity to the region to which it belongs. The membership values of the pixels are determined using Gamma distribution as described in next section.

5.4 Gamma distribution

The general formula for the probability density function of Gamma distribution is

$$f(x) = \frac{\left(\frac{(x-v)}{\beta}\right)^{\gamma-1} \exp\left(-\frac{(x-v)}{\beta}\right)}{\Gamma(\gamma)} \quad x \geq v; \gamma, \beta > 0$$

(11)

Where γ is shape parameter, v is location parameter, β is the scale parameter and Γ is the Gamma function given by

$$\Gamma(\gamma) = \int\limits_{0}^{\infty} u^{\gamma-1} e^{-u} du$$

We have the following cases:

Case 1: When $v = 0$ and $\beta = 1$, the distribution assumes the form

$$f(x) = \frac{x^{\gamma-1} \exp(-x)}{\Gamma(\gamma)} \quad x \geq 0; \quad \gamma > 0$$

(12)

which is known as the standard Gamma distribution.

Case 2: When $v \neq 0$, $\beta = 1$ and $\gamma = 1$, the Gamma distribution takes the form

$$f(x) = \exp(-(x-v))$$

(13)

Replacing v in equation (13) by μ_0 and μ_1 separately from equations (9) and (10), the membership function for the background and object becomes

$$\mu(f_{ij}) = \exp(-c.|f_{ij} - \mu_0|) \ if \ f_{ij} \leq t,$$ for background

$$= \exp(-c.|f_{ij} - \mu_1|) \ if \ f_{ij} > t,$$ for object

(14)

where t is any chosen threshold as stated above. It may be pointed out that in the membership function, the constant 'c' is taken to ensure membership of the gray level feasible in the range $[0,1]$. Here 'c' is chosen as

$$c = \frac{1}{(f_{max} - f_{min})}$$, where f_{min} and f_{max} are the minimum and maximum gray level in the image respectively. The absolute value of the distance between the mean of the region to which a pixel belongs and the gray level of that pixel is considered.

For tri-level thresholding, where there are three regions in the image, two thresholds values t_1 and t_2 are selected such that $0 \leq t_1 \leq t_2 \leq L-1$, where L is the maximum gray level of the image. Extending the concept of bi-level thresholding, the membership function in case of tri-level thresholding will take the form

$$\mu(f_{ij}) = \exp(-c_1.|f_{ij} - \mu_0|) \ \ if \ \ f_{ij} \le t_1$$

$$= \exp(-c_1.|f_{ij} - \mu_1|) \ \ if \ \ t_1 < f_{ij} \le t_2$$

$$= \exp(-c_1.|f_{ij} - \mu_1|) \ \ if \ \ f_{ij} > t_2 \tag{15}$$

where, μ_0, μ_1 and μ_2 are the average gray levels for the three regions separated by the thresholds t_1 and t_2 and the constant ' c_1 ' is like ' c ' in Eq. (14).

5.5 Fuzzy divergence between two images

Let $\mu_A(f_{ij})$ and $\mu_B(f_{ij})$ be the membership values of the pixels in the image and f_{ij} is $(i, j)th$ pixel in image A. Then in view of equation (8) fuzzy divergence between A and B is given by

$$M^F_*(A,B) =$$

$$\sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \frac{(\mu_A(f_{ij}) - \mu_B(f_{ij}))^2}{2} \left[\frac{1}{\mu_A(f_{ij}) + \mu_B(x_i)} + \frac{1}{2 - \mu_A(f_{ij}) - \mu_B(f_{ij})} \right] \tag{16}$$

5.6 Methodology

For bi-level or multilevel thresholding a searching methodology based on image histogram is employed here. The region between the two successive peaks is the region for searching. If there are more valleys (in case of multimodal histogram) succeeding and preceding peaks of each valley are noted and accordingly the search regions are selected. For unimodal histogram, linear search is employed for selecting the threshold. For each threshold, the membership values of all the pixels in the image are found out using the above procedure. For each threshold value, the membership values of the thresholded image are compared with an ideally thresholded image. Thus equation (16) reduces to

$$M^F_*(A,B) =$$

$$\sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \frac{(\mu_A(f_{ij}) - 1)^2}{2} \left[\frac{1}{\mu_A(f_{ij}) + 1} + \frac{1}{2 - \mu_A(f_{ij}) - 1} \right]$$

$$= \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \frac{(\mu_A(f_{ij}) - 1)^2}{2} \left[\frac{1}{\mu_A(f_{ij}) + 1} + \frac{1}{1 - \mu_A(f_{ij})} \right]$$

$$= \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \left[\frac{1 - \mu_A(f_{ij})}{1 + \mu_A(f_{ij})} \right] \tag{17}$$

As the membership values of each pixel in an identically threshold image (which is image B in this case) are taken as unity.

An ideally thresholded image is that image which is precisely segmented so that the pixels, which are in the object or the background region, belong totally to the respective regions. From the divergence value of each pixel between the ideally segmented image and the above chosen thresholded image, the fuzzy divergence is found out. It is expected that the membership values of each pixel in the thresholded image should lie close to that of the ideally thresholded image for good thresholding. If a pixel lies in the object/background region, it should contribute more to the corresponding object/background region to which it belongs. In this way, for each threshold, divergence of each pixel is determined according to Eq. (17) and the cumulative divergence is computed for the whole image. The minimum divergence is selected and the corresponding gray level is chosen as the optimum threshold. Here the minimum divergence yields a measure of the maximum belongingness of each object pixel to the object region and that of each background pixel to the background region. After thresholding, the thresholded image leads almost towards the ideally thresholded image.

5.7 Experimental result

The thresholding algorithm described above is tested on unimodal, bimodal and multimodal images. In order to evaluate the effectiveness of the proposed method, severel images were tested. Here we present experimental result of a bimodal image in png format. Let input image 'Coins' of size 128×128 (bimodal)'

Fig. 1 and Fig. 2 show a 'coins' image of size 128×128 and its bimodal histogram. Fig. 3 shows the thresholded image, which is thresholded at gray level 96.
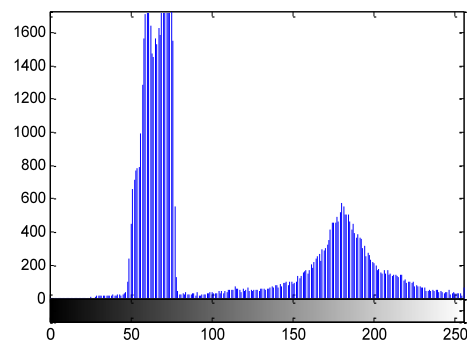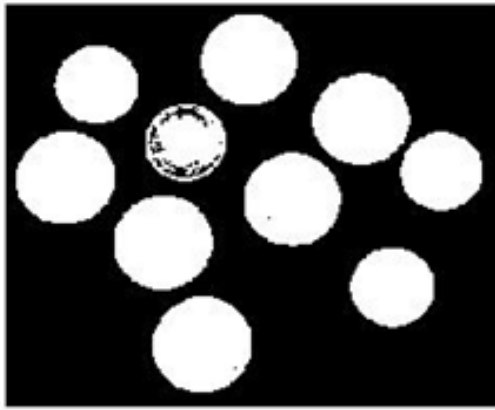

Fig. 1

Fig. 2

Fig. 3

VI. Conclusion

In this communication an approach to develop measures of fuzzy directed divergence using aggregation operators is proposed. The proposed measure is not a distance measure but there is possibility of development of distance measures. To add flexibility in applications the divergence (distance) measures may be generalized by using a parameter. In the literature related to image segmentation, the optimum threshold is obtained either by maximizing the fuzzy entropy or by minimizing the fuzzy divergence. Here the optimum threshold is obtained by minimizing the proposed fuzzy divergence. The comparison of proposed measure with the existing measures of fuzzy directed divergence in context of image segmentation is not done, but this is a measure to its own right and can be used for thresholding in some situations because different measures have their suitability in different situations.

References

[1] Shannon, C.E., The Mathematical Theory of Communications. Bell Syst. Tech. Journal, 1948, 27: 423–467.

[2] Renyi, A.,On Measures of Entropy and Information. Proc. 4th Berkeley Symp. Math. Stat Probab.1, 1961 :547–561.

[3] Kullback,S., R.A. Leibler, On Information and Sufficiency. Ann. Math. Statist., 1951,22 : 79-86.

[4] Taneja,I.J.,Generalized Information Measures and their Applications - On-line book :http: //www. mtm.ufsc.br/~ taneja/book/book.html, 2001.

[5] Basseville, M., Divergence measures for statistical data processing.Publications Internes de l'IRISA ,Novembre 2010.

[6] Esteban ,Marí Dolores, Domingo Morales,,A summary on entropy statistic,s"Kybernetika,1995, 31(4): 337–346.

[7] Zadeh,L.A., Fuzzy Sets, Information and control, 1965, 8 : 338 - 353.

[8] De Luca,A., Termini,S., A definition of non-probabilistic entropy in the setting of fuzzy set theory, Inform. and Control, 1971, 20: 301 - 312.

[9] Bhandari, D., Pal,N. R., Some new information measures for fuzzy sets, Information Sciences, 1993, 67: 204 – 228

[10] Bhatia,P.K., Singh,S., On some divergence measures between fuzzy sets and aggregation operations (Communicated).

[11] Otsu, K., A threshold selection method from gray level histograms. IEEE Trans. Systems Man Cyber.,1979, 9: 62-66.

[12] Kapur, J.N., Sahoo, P.K., Wong, A.K.C., A new method of gray level picture thresholding using the entropy of the histogram. Comput. Vision Graphics Image Process, 1985, 29: 273-285.

[13] Sahoo, P.K., Soltani, S., Wong, A.K.C., Chen, Y.C., A survey of thresholding techniques. Comput. Vision Graphics Image Process.,1988: 233-260.

[14] Sahoo, P.K., Wilkins, C., Yager, Threshold selection using Renyi's entropy. Pattern Recognition, 1997, 30: 71-84.

[15] Sahoo, P.K, Wong, A.K.C., A gray level threshold selection method based on maximum entropy principle. IEEE Ttans. Man Systems and Cyber., 1989, 19 : 866-871.

[16] Brink, A.D., Pendcock, N.E.,Minimum cross-entropy threshold selection. Pattern Recognition, 1996, 29: 179-188.

[17] Pal, S.K., Dasgupta, A., Spectral fuzzy sets and soft thresholding. Inform. Sci,1992,. 65 : 65-97.

[18] Huang, L.K., Wang, M.J., Image thresholding by minimizing the measure of fuzziness. Pattern Recognition, 1995, 28 (1) : 41-51.

[19] Ramar, K. et al., Quantitative fuzzy measures for threshold selection. Pattern Recognition Lett..2000, 21 : 1-7.

[20] Cheng, H.D., Chen, H.H., Image segmentation using fuzzy homogeneity criterion. Inform. Sci., 1997, 98: 237-262.

[21] Chaira, T., Ray, A.K., Segmentation using fuzzy divergence Pattern Recognition Letters , 2003, 24: 1837-1844.

[22] Couso, I., Janis,V.,Montes,S., Fuzzy Divergence Measures, Acta Univ. M. Belii Math., 2004, 8: 21 – 26.

[23] Qing, M., Li, T., Some Properties and New Formulae of Fuzzy Entropy, Proceedings of the 2004 IEEE International Conference on Networking, Sensing & Control, Taipei, Taiwan, March 21-23, 2004: 401-406.

[24] Bhatia,P.K.,Singh,S.,Kumar,V., Some New Divergence Measures and Their Properties, Int. J. of Mathematical Sciences and Applications, 2011, 1(3): 1349 - 1356.

[25] Taneja,I.J., On Mean divergence measures. available at arXiv:math/0501297v1[math.ST] 19 Jan 2005.

[26] Taxt,T.,Flynn,P. J. and Jain,A. K., Segmentation of document images, IEEE Trans. Pattern Analysis Mach. Intell.,1989, 11 (12): 1322-1329

Authors' Profiles

P.K Bhatia, Professor of Mathematics at the DCR University of Science & Technology, Murthal (Sonepat), India. He has taught undergraduate and post graduate students for over twenty five years. He has been supervising Ph.D. students. He has published a number of research papers in journals of national and international repute. His areas of interest are "Mathematical Information Theory" & "Reliability Modeling".

Surender Singh, Assistant Professor at School of Mathematics, Shri Mata Vaishno Devi University, Katra (J & K), India. He has taught undergraduate and post graduate students for over ten years. He is also working towards his Ph.D under supervision of Prof. P.K Bhatia. His research directions include Information, divergence measures and their applications.