Computers and Mathematics with Applications 60 (2010) 2426–2432

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

A novel automatic microcalcification detection technique using Tsallis entropy & a type II fuzzy index

Mohanalin *, Beenamol, Prem Kumar Kalra, Nirmal Kumar

Department of Electrical Engineering, IIT Kanpur, UP-208016, India

ARTICLE INFO

Article history:
Received 18 August 2009
Received in revised form 12 August 2010
Accepted 12 August 2010

Keywords:
Tsallis entropy
Type II fuzzy index
Shannon entropy
Mammograms
Microcalcification

ABSTRACT

This article investigates a novel automatic microcalcification detection method using a type II fuzzy index. The thresholding is performed using the Tsallis entropy characterized by another parameter '*q*', which depends on the non-extensiveness of a mammogram. In previous studies, '*q*' was calculated using the histogram distribution, which can lead to erroneous results when pectoral muscles are included. In this study, we have used a type II fuzzy index to find the optimal value of '*q*'. The proposed approach has been tested on several mammograms. The results suggest that the proposed Tsallis entropy approach outperforms the two-dimensional non-fuzzy approach and the conventional Shannon entropy partition approach. Moreover, our thresholding technique is completely automatic, unlike the methods of previous related works. Without Tsallis entropy enhancement, detection of microcalcifications is meager: 80.21% Tps (true positives) with 8.1 Fps (false positives), whereas upon introduction of the Tsallis entropy, the results surge to 96.55% Tps with 0.4 Fps.

1. Introduction

Several articleshave been published highlighting the challenges existing in microcalcification cluster (Mcs) detection for early breast cancer diagnosis. Lately, Tsallis entropy (TE) based works have created a lot of interest [1]. It is proven that TE gives better thresholding results [2]. [3] compared the performance of traditionally used SE with TE and concluded that TE is far superior in detecting Mcs in mammograms. But there a histogram distribution technique was employed for calculating '*q*', which can lead to erroneous results when pectoral muscles are included. In the present study, a new technique based on type II fuzzy theory is proposed for calculating '*q*' optimally. The proposed approach has been tested on various images, and the results have demonstrated that the proposed TE approach outperforms the two-dimensional non-fuzzy approach and conventional SE partition approach.

The paper is organized as follows. Section 2 deals with the basics of TE and the proposed algorithm. Section 3 demonstrates the role of the type II fuzzy index in '*q*' identification. Sections 4 and 5 discuss the implementation of enhancement and the detection of Mcs respectively. The validation procedure adopted and the conclusions from our experiment are presented in Section 6.

2. The proposed algorithm

This section proposes TE based detection of Mcs using a type II fuzzy set. The proposed algorithm is shown in Fig. 1. The mammogram can be looked at as three individual objects fused together, namely, the picture background, the tissue background (including the fatty area) and Mcs (ROI). Generally in medical images the background remains black. Hence,

* Corresponding author.
E-mail address: mohanlee@iitk.ac.in (Mohanalin).

0898-1221/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2010.08.038

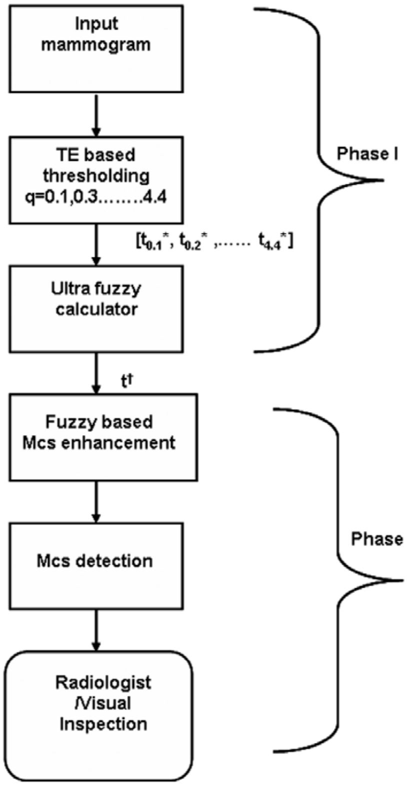

Fig. 1. Proposed algorithm using TE for detection of Mcs.

the intensity level of the background must be less than the average value of the intensity of the image. This object is of no significance and can be filtered off by employing the mean. The mean value 'k' is calculated using Eq. (1):

$$k = \frac{1}{D} * \sum^{X} \sum^{Y}_{m,n\in G} g_{m,n} \tag{1}$$

with the following key:

M, N — the dimensions of the image;
G — the intensities larger than 100 for normal images and 10 for denser images;
D — the number of pixels;
g_{mn} — the grey level at coordinates m and n.

The Mcs will not be affected during this process, as generally the intensity of the Mcs is higher than the average of the image.

The probability distribution of grey levels in the mammogram can be written as $p_1, p_{k+1}, \ldots, p_N$. The tissue background (A) of the mammogram can be formulated as

$$\frac{p_k}{p_t - P_{k-1}}, \frac{p_{k+1}}{P_t - P_{k-1}}, \ldots, \frac{P_t}{P_t - P_{k-1}}.$$

The region of interest, i.e. the Mcs, in the mammogram (B) can be framed in terms of equations as follows:

$$\frac{p_{t+1}}{1-P_t}, \frac{p_{t+2}}{1-P_t}, \ldots, \frac{p_N}{1-P_t}$$

$$P_{k-1} = \sum_{i=1}^{k-1} p_i$$

$$P_t = \sum_{j=k}^{t} p_j.$$