Quantification and Segmentation of Breast Cancer Diagnosis: Efficient Hardware Accelerator Approach

KHAIRULNIZAM OTHMAN
Central of Diploma Studies,
University Tun Hussein Onn 84600, Johor,
MALAYSIA
nizam@uthm.edu.my   http://fkee.uthm.edu.my/mint-src/

AFANDI AHMAD
Embedded Computing Research Cluster
Microelectronics and Nanotechnology - Shamsuddin Research Centre (MiNT-SRC)
Faculty of Electrical and Electronic Engineering
University Tun Hussein Onn 84600, Johor,
MALAYSIA
afandia@uthm.edu.my

Abstract: - This research presents VLSI architecture for image segmentation. The architecture is based on the fuzzy c-means algorithm with spatial constraint for reducing the misclassification rate. In the architecture, the usual iterative operations for updating the homogeneus membership matrix and cluster centroid are merged into one single updating process to evade the large storage requirement. In addition, an efficient pipelined circuit is used for the updating process for accelerating the computational speed. Experimental results show that the proposed circuit is an effective alternative for real-time image segmentation with low area cost (time) and low misclassification rate.

Key-Words: - fuzzy c-means, image segmentation, clustering, FPGA, chip

1 Introduction

Breast cancer is the commonest cancer among Malaysian women in all ethnic groups [1]. A woman in Malaysia has 1 in 20 chance of getting a breast cancer in her lifetime. According to the National Cancer Registry, the Age Standardized Rate (ASR) of female breast cancer in Malaysia is 47.4 per 100,000 populations in 2003 to 2005 and this dropped to 39.3 in 2006. However, the ASR is higher than those in other Asian countries (Beijing, 24.6, Hiroshima 36.6, Chennai 23.9 and Seoul 20.8) [1]. There are variations in the incidence rates of breast cancer among the three main ethnic groups in Malaysia. Amongst the Chinese, the ASR is highest at 59.9 per 100,000 population, for the Indians, the ASR is 54.2 per 100,000 and it is lowest in the Malays at 34.9 per 100,000 population. Breast cancer accounted for 31.3% of the total number of the new cases in women [1]. These differences may be attributed to the differences in lifestyle, diet and reproductive behavior namely those that relates to childbearing and breastfeeding practices.

Studies on breast cancer have been approached for many years from different angles of importance; be it the cause of the disease, the detection of the disease, the diagnosis systems, method of treatments before and after surgery. These studies have been divided into two paradigms, one which defines breast cancer disease as a local and regional disease; and another as a systematic disease. The role of a surgeon alone is adequate for the breast cancer that is initially at a local and regional disease process. However, when the breast cancer is a systematic disease then the involvement of systematic treatments is needed besides the surgery alone to cure it and there will be a role for earlier detection of breast cancer.

Breast cancer screening programs attempt to detect and eradicate cancer at the earliest possible stage to increase the rate of survival amongst women. The early detection of breast cancer greatly improves the prognosis. One of the earliest signs of cancer is the formation of clusters of micro-calcifications.

Although clinicians have in the past been reluctant to use computational approaches in diagnostic applications, it is now considered that the use of advanced high performance methodologies will be useful in aiding clinicians in diagnosis and treatment planning. At present, segmentations are mainly consulted manually by medical specialists. Although this task appears simple, the reality is that an in-depth knowledge of anatomy and physiology is required. Essentially, the expert observes a particular image, determines borders between regions and classifies each region.

In addition to this, to identify smaller image features, contrast modifications are often needed. Although, for a typical data, the entire expert manual segmentation can take several hours to compute, this approach is perhaps the most reliable and accurate method of medical image segmentation. This is due to the immense complexity of the human visual system, a system well suited to this task. This rise an opportunity for us to develop hardware for prospectively evaluated a computer-aided detection (CAD) device used with data from mammography by assessing breast cancers detected; tumor sizes, histology, and stage. Insight acquired through this research is expected to be useful for the development hardware, which is identified breast cancer to be relevant to bio-medic engineering problem solving.

Given the massive volumes of mammography image data to be processed, and the considerable computational complexity of the algorithms that are involved, it is clear that scheme can be put to efficient practical use only if efficient hardware co-processing strategies are developed. The appearance of fast reconfigurable field programmable gate arrays (FPGAs) brings about a new path for the design of such systems. FPGAs can perform mathematical operations on an entire vector or matrix at the same time, and the current generation of DSP-capable FPGAs yields ultra-high performance and highly flexible signal-processing systems. FPGA-based implementation also has the advantage of short design period and low cost of fabrication, which are suitable for partial or complete acceleration of various algorithms. These factors make FPGAs an attractive platform for implementing the proposed segmentation system. However, the FPGAs present limited resources to implement complex hardware. Hence multiple FPGA board solutions or run-time context switching techniques provide the possibility of implementing such complex system by using time multiplexing

strategies. To ensure system integration and tight coupling between software processing and hardware acceleration using FPGAs, it is essential to develop suitable host applications that can be used for the co-design process.

Image segmentation plays an important role in computer vision and image analysis. The segmentation results can be used to identify regions of interest and objects in the scene, which is very beneficial to the subsequent image analysis or annotation. The fuzzy c-means algorithm (FCM) [2] is one of the most-used techniques for image segmentation. The accuracy of FCM is due to the employment of fuzziness for the clustering of each image pixel. This enables the fuzzy clustering methods to retain more information from the original image than the crisp or hard segmentation. Although the original intensity-based FCM algorithm functions well on segmenting most noise-free images, it fails to segment images corrupted by noise, outliers and other medical imaging. The FCM with spatial constraint (FCM-S) algorithms [3] – [5] have been proposed to solve this problem by incorporating spatial information into original FCM objective function. However, as compared with the original FCM algorithm, the FCM-S algorithms have higher computational complexities for membership coefficients computation and centroid updating. In addition, similar to the original FCM algorithm, the size of homogeneus membership matrix grows as the product of data set size and number of classes in the FCM-S. As a result, the corresponding memory requirement may prevent the algorithm from being applied to images with high dimension.

To accelerate the computational speed and reduce the memory requirement of the original FCM, a number of algorithms [6] – [9] have been proposed. These fast algorithms can be extended for the implementation of FCM-S. However, most of these algorithms are implemented by software, and only moderate acceleration can be achieved. In [10] – [12], hardware implementations of FCM are proposed. Nevertheless, the design in [9] is based on analog circuits. The clustering results therefore are difficult to be directly used for digital applications. Although the architecture shown in [11] adopts digital circuits, the architecture aims for applications with only two classes. In addition, it may be difficult to extend the architecture for the hardware implementation of FCM-S. The architecture presented in [12] operates with only a fixed degree of fuzziness $m = 2$ for the original FCM. The

flexibility for selecting other degrees of fuzziness may be desired to further improve the FCM performance. In addition, similar to [11], the architecture presented in [12] cannot be directly used for the hardware implementation of FCM-S.

The architecture relaxes the restriction on the degree of fuzziness. The relaxation requires the employment of n-th root and division operations for membership coefficents and centroid computation. A pipeline implementation for the FCM-S therefore may be difficult. To solve the problem, in the proposed architecture, the n-th root operators and dividers are based on simple table lookup, multiplication and shift operations. Efficient pipeline circuits can then be adopted to enhance the throughput for fuzzy clustering.

To reduce large memory size for storing membership matrix, the proposed architecture combines the usual iterative updating processes of membership matrix and cluster centroid into a single updating process. In the architecture, the updating process is separated into three steps: pre-computation, membership coefficents updating, and centroid updating. The pre-computing step is used to compute and store information common to the updating of different membership coefficents. This step is beneficial for reducing the computational complexity for the updating of membership coefficents. The membership updating step computes new membership coefficents based on a fixed set of centroids and the results of the pre-computation step. All the membership coefficents associated with a data point will be computed in parallel in this step. The computation time of the FCM-S therefore will be effectively expedited.

The centroid updating step computes the centroid of clusters using the current results obtained from the membership updating step. The weighted sum of data points and the sum of membership coefficents are updated incrementally here for the centroid computation. This incremental updating scheme eliminates the requirement for storing the entire membership coefficents.

The proposed architecture has been implemented on field programmable gate array (FPGA) devices [13] so that it can operate in conjunction with a soft-core CPU [14]. Using the reconfigurable hardware, we are then able to construct a system on programmable chip (SOPC) system for image segmentation. The proposed architecture attain lower classification

error rate in the presence of noise. In addition, compared with its software counterpart running on the 2.53 GHz Intel Core 2 Duo, our system has significantly lower computational time. All these facts demonstrate the effectiveness of the proposed architecture.

2 Contributions to the Development of c-Means Clustering Models Introduction

We first give a brief review of the FCM algorithm. Let $X = \{x_1, ..., x_t\}$ be a data set to be clustered by the FCM algorithm into c classes, where t is the number of data points in the design set. Each class i, $1 \le i \le c$, is characterized by its centroid v_i. The goal of FCM is to minimize the following cost function:

$$J = \sum_{i=1}^{c} \sum_{k=1}^{t} \mu_{ik}^{m} \left\| p_i - v_k \right\|^2$$

$$(1)$$

where $u_{i,k}$ is the membership of x_k in class i, and $m > 1$ indicates the degree of fuzziness. The cost function J is minimized by a two-step iteration in the FCM. In the first step, the centroids $v_1, ..., v_c$, are fixed, and the optimal membership matrix $\{u_{i,k}, i = 1, ..., c, k = 1, ..., t\}$ is computed by

$$u_{i,k} = \left(\sum_{j=1}^{c} \left(\frac{\left\| x_k - v_i \right\|}{\left\| x_k - v_j \right\|} \right)^{2/(m-1)} \right)^{-1}$$

$$(2)$$

After the first step, the membership matrix is then fixed, and the new centroid of each class i is obtained by

$$v_i = \frac{\left(\sum_{k=1}^{t} u_{i,k}^{m} x_k \right)}{\left(\sum_{k=1}^{t} u_{i,k}^{m} \right)}$$

$$(3)$$

A variant of FCM for image segmentation is FCM-S, whose objective function is [2]

$$J = \sum_{i=1}^{c} \sum_{k=1}^{t} v_{i,k}^{m} \|x_k - v_k\|^2 + \frac{\alpha}{card(\Gamma)} \sum_{i=1}^{c} \sum_{k=1}^{t} u_{i,k}^{m} \sum_{j\in\Gamma} \|x_j - v_i\|^2$$

(4)

where Γ is the set of neighbours associated with x_k, and the $Card(\Gamma)$ is the cardinality of the set Γ. The parameter α determines the degree of penalty. The necessary conditions locally minimizing J are then given by

$$u_{i,k} = \frac{\left(\|x_k - v_i\|^2 + \frac{\alpha}{card(\Gamma)} \sum_{j\in\Gamma} \|x_j - v_i\|^2 \right)^{-1/(m-1)}}{\sum_{n=1}^{c} \left(\|x_k - v_n\|^2 + \frac{\alpha}{card(\Gamma)} \sum_{j\in\Gamma} \|x_j - v_n\|^2 \right)^{-1/(m-1)}}$$

(5)

$$v_i = \frac{\sum_{k=1}^{t} u_{i,k}^{m} \left(x_k + \frac{\alpha}{card(\Gamma)} \sum_{j\in\Gamma} x_j \right)}{(1+\alpha) \sum_{k=1}^{t} u_{i,k}^{m}}$$

(6)

The disadvantages of Equations (5) and (6) are the high computational complexities for computing $u_{i,j}$ and v_i. To accelerate the computation, observe from (3) that by simple manipulation,

$$\frac{1}{card(\Gamma)} \sum_{j\in\Gamma} \|x_j - v_i\|^2$$ can be equivalently written as

$$\frac{1}{card(\Gamma)} \sum_{j\in\Gamma} \|x_j - v_i\|^2 = \left(\frac{1}{card(\Gamma)} \sum_{j\in\Gamma} \|x_j - \bar{x}_k\|^2 + \|\bar{x}_k - v_i\|^2 \right)$$

(7)

Where

$$\bar{x}_k = \frac{1}{card(\Gamma)} \sum_{j\in\Gamma} x_j$$

(8)

Note that $\bar{x}_k$ can be computed in advance, and the minimization of J in Equation (4) is equivalent to the minimization of the following cost function.

$$J = \sum_{i=1}^{c} \sum_{k=1}^{t} u_{i,k}^{m} \|x_k - v_i\|^2 + \alpha \sum_{i=1}^{c} \sum_{k=1}^{t} u_{i,k}^{m} \|\bar{x}_k - v_i\|^2$$

(9)

Necessary conditions on $u_{i,j}$ and v_i for locally minimizing J can be derived are follows.

$$u_{i,k} = \frac{\left(\|x_k - v_i\|^2 + \alpha \|\bar{x}_j - v_i\|^2 \right)^{-1/(m-1)}}{\sum_{j=1}^{c} \left(\|x_k - v_j\|^2 + \alpha \|\bar{x}_k - v_j\|^2 \right)^{-1/(m-1)}}$$

(10)

$$v_i = \frac{\sum_{k=1}^{t} u_{i,k}^{m} \left(x_k + \alpha \bar{x}_k \right)}{(1-\alpha) \sum_{k=1}^{t} u_{i,k}^{m}}$$

(11)

The FCM and FCM-S algorithm requires large number of floating point operations. Moreover, from Equations [15], [16], [17] and [18], it follows that the membership matrix needs to be stored for the computation of cost function and centroids. As the size of the membership matrix grows with the product of t and c, the storage size required for the FCM may be impractically large when the data set size and/or the number of classes become high.

The pre-computation unit is used for reducing the computational complexity of the membership coefficients calculation. Observe that $u_{i,k}$ in Equation (2) can be rewritten as

$$u_{i,k} = \left\| x_k - v_i \right\|^{-2/(m-1)} p_k^{-1}$$

(12)

Where

$$p_k = \sum_{j=1}^{c} \left(\frac{1}{\left\| x_k - v_j \right\|^2} \right)^{1/(m-1)}$$

(13)

Given x_k and centroids v_1, ..., v_c, membership coefficients $u_{1,k}$, ..., $u_{c,k}$ have the same P_k. Therefore, the complexity for computing membership coefficients can be reduced by calculating P_k in the pre-computation unit. Without loss of generality, the degree of fuzziness m can be expressed as

$$m = a/b$$

(14)

where both a and b are integers. Because m should be larger than 1, it follows that $a > b > 0$. Let

$$r = b, n = a - b$$

(15)

We then can rewrite Equation (13) as

$$p_k = \sum_{j=1}^{c} \left(\left\| x_k - v_j \right\| \right)^{-2r/n}$$

(16)

Based on Equation (16), we see that the n-th root operation is required for the implementation of p_k. In the proposed architecture, a novel n-th root circuit is adopted so that P_k can be implemented in a pipelined fashion. In the proposed n-th root circuit, the goal is to compute $\sqrt[n]{Y}$, where

$$Y = 1 + 2^{-1} y_1 + 2^{-2} y_2 + ... + 2^{-(2q-1)} y_{2q-1}$$

That is, Y is a $2q$-bits real number such that $1 < Y < 2$. We separate Y into two portions Y_h and Y_l as shown below

$$Y_h = 1 + 2^{-1} y_1 + 2^{-2} y_2 + ... + 2^{-(2q-1)} y_{q-1}$$

(17)

$$Y_l = 2^{-(q+1)} y_{q+1} + 2^{-(q+2)} y_{q+2} + ... + 2^{-(2q-1)} y_{2q-1}$$

(18)

For the sake of simplicity, we first consider the computation of $\sqrt{Y}$. Observe that

$$\sqrt{Y} = \frac{Y}{(Y_h + Y_l)^{1/2}} = \frac{Y/Y_h^{1/2}}{(1 + Y_l/Y_h)^{1/2}} = \frac{Y}{Y_h^{1/2}} \left(1 - \frac{Y_l}{2Y_h} + \frac{3Y_l^2}{8Y_h^2} ...\right)$$

By retaining the first two terms of the Taylor series, $\sqrt{Y}$ can be approximated by

$$\sqrt{Y} \approx \frac{Y}{Y_h^{1/2}} \left(1 - \frac{Y_l}{2Y_h}\right) = \frac{Y(Y_h - Y_l/2)}{Y_h^{3/2}}$$

From Equations (17) and (18), we conclude that $Y_h > 2^q Y_l$. Therefore, the maximum error of the approximation is less than 2^{-2q}. Following the same procedure, it can also be found that

$$\sqrt[3]{Y} = \frac{Y}{(Y_h + Y_l)^{2/3}}$$

$$= \frac{Y/Y_h^{2/3}}{(1 + Y_l/Y_h)^{2/3}}$$

$$= \frac{Y}{Y_h^{2/3}}(1 - \frac{2Y_l}{3Y_h} + ...) \approx \frac{Y(Y_h - 2Y_l/3)}{Y_h^{5/3}}$$

These results can be extended for any $n \geq 2$ as follows:

$$\sqrt[n]{Y} \approx \frac{Y(Y_h - (n-1)Y_l/n)}{Y_h^{(2n-1)/n}}$$

(19)

The n-th root circuit based on Equation (19) is shown in Figure 2, which consists of two tables, two multipliers, and one adder. The tables store $(n-1)Y_l/n$ and $Y_h^{(2n-1)/n}$ for all the possible values of Y_l and Y_h. Although it is possible to construct a table directly for $\sqrt[n]{Y}$, the number of entries in the table would be 2^{2q-1} because Y contains $2q$ bits. By contrast, both Y_h and Y_l consist of only q bits. The number of entries in each table shown in Figure 2 is only 2^{q-1}. Consequently, the proposed circuit is able to perform fast and accurate computation while maintaining low area cost.
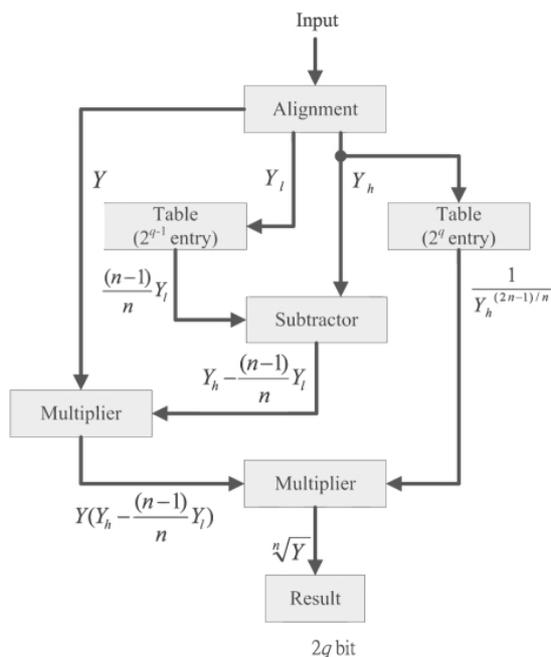
Fig. 1: The architecture of n-th root unit.

3.0   The Proposed Architecture

The goal of the proposed architecture is to implement the FCM-S algorithm in hardware. The architecture is based on a novel pipeline circuit to provide high throughput for fuzzy clustering. It is also able to eliminate the requirement for storing the large membership matrix for the computation of cost function and centroids.

3.1.   FCM-S Architecture

The goal of the mean computation unit is to evaluate the mean value $\bar{x}_k$ defined in Equation (8). The main architecture of FCM-S is the fuzzy clustering unit, which computes the membership coefficients and centroids of FCM-S. Therefore, our discussion in this subsection will focus on the fuzzy clustering unit of the FCM-S. Using Equations (14) and (15), we can rewrite the membership coefficients of FCM-S defined in Equation (10) as

$$u_{i,k}^m = ((\|x_k - v_i\|^2 + \alpha\|\bar{x}_k - v_i\|^2)^{1/n} p_k^{1/r})^{-(n+r)}$$

(23)

Where

$$p_k = \sum_{j=1}^{c} (\|x_k - v_j\|^2 + \alpha\|x_k - v_j\|^2)^{-r/n}$$

(24)

Similar to the original FCM, it follows from Equation (24) that the computation of P_k can also be separated into c terms, where the j-th term involves the computation of $(\|x_k - v_j\|^2 + \alpha\|\bar{x}_k - v_j\|^2)^{-r/n}$. Figure 9 shows the architecture for the computation of each $(\|x_k - v_j\|^2 + \alpha\|\bar{x}_k - v_j\|^2)^{-r/n}$. From Figure 9, we see that the architecture can also be implemented as a 4-stage pipeline, similar to that shown in Figure 3(b) for computing $(\|x_k - v_j\|)^{-2r/n}$. Therefore, the pre-computation unit for FCM-S can be realized as a $4c$ stage pipeline.
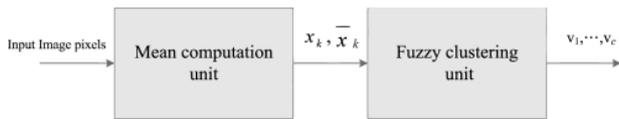
Fig. 2: The FCM-S architecture.
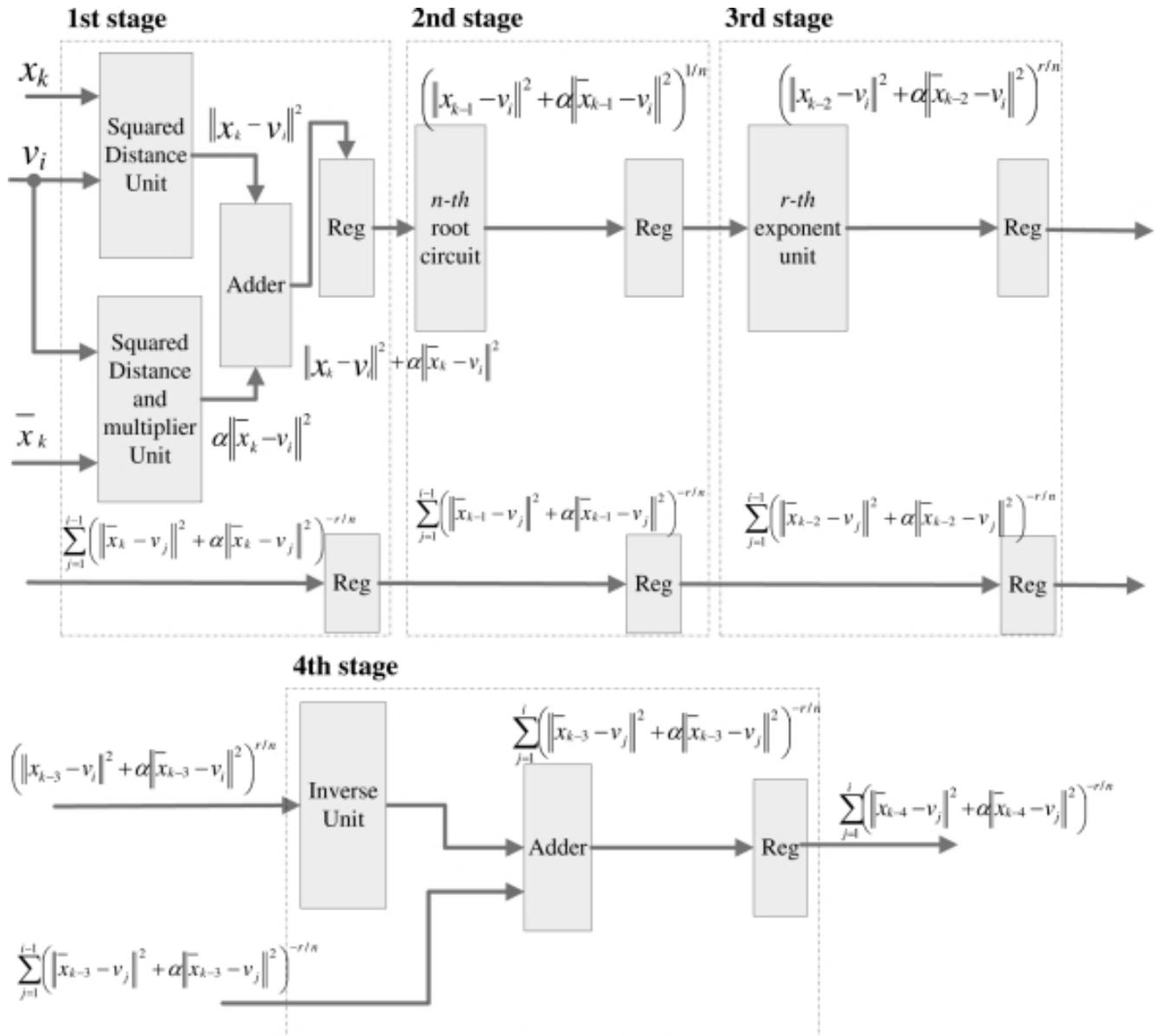
Fig. 3: The circuit for evaluating $(\left\| x_k - v_j \right\|^2 + \alpha \left\| \overline{x_k - v_j} \right\|^2)^{-r/n}$

The only difference is that the first stage of the pipeline in Figure 3 has higher area and computational complexities. There are two squared distance calculation units and one adder at the first stage of the pipeline in Figure 3.

Figures 4 – 6 depict the architecture for membership coefficients updating, centroids updating and cost function computation for FCM-S based on Equations (9), (11) and (23), respectively. Similar to the original FCM algorithm, the proposed FCM-S architecture computes the centroids and cost

function incrementally. In the FCM-S, the incremental centroid for the i-th cluster up to data point x_k is defined as

$$v_i(k) = \left(\frac{\sum_{n=1}^{k} u_{i,n}^m (x_n + \alpha \bar{x}_n)}{(1+\alpha) \left(\sum_{n=1}^{k} u_{i,n}^m \right)} \right)$$

(25)

$$J(k) = \sum_{i=1}^{c} \sum_{n=1}^{k} u_{i,n}^m \left(\|x_n - v_i\|^2 + \alpha \|\bar{x}_k - v_j\|^2 \right)$$

(26)

As shown in Figures 5 and 6, the goals of the centroids updating unit and the cost function computation unit are to compute $v_i(k)$ and $J(k)$, respectively. As $k = t$, the $v_k(i)$ and $J(k)$ in Equations (25) and (6) will becomes $v(i)$ in Equation (11) and J in Equation (9), respectively.
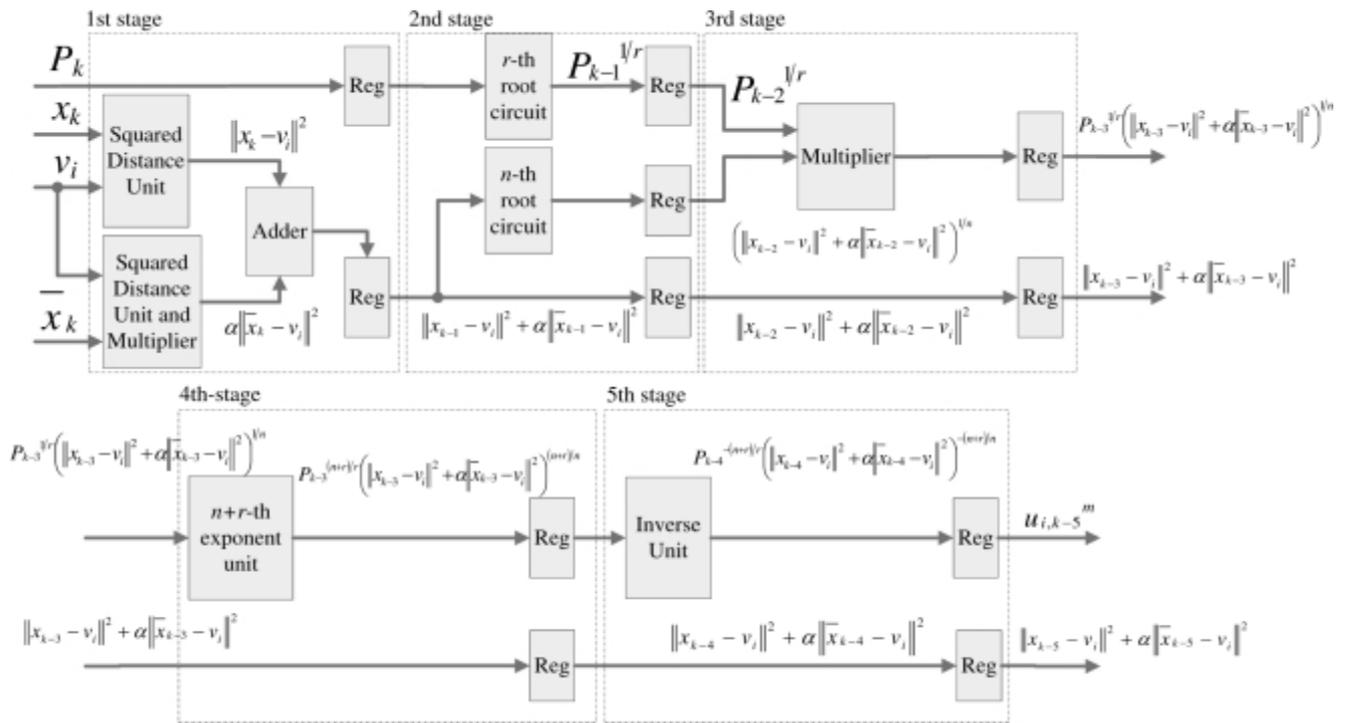
Fig. 4: The circuit for evaluating $u_{i,k}^m$ for FCM-S.
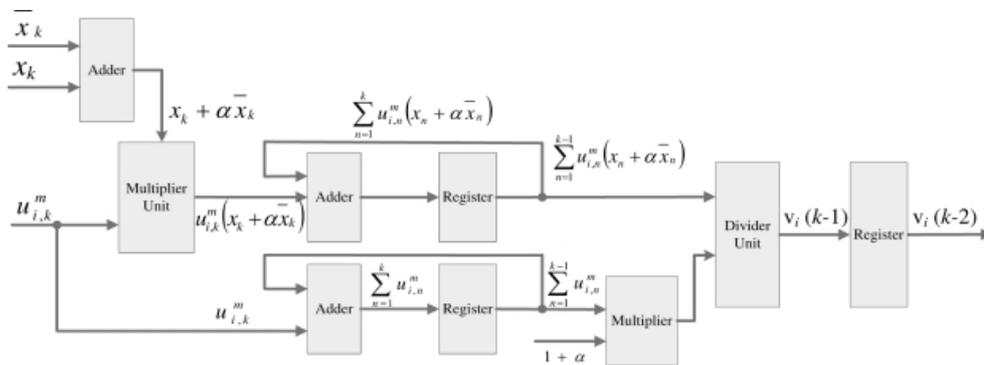
Fig. 5: The circuit for calculating $v_i(k)$ for FCM-S.
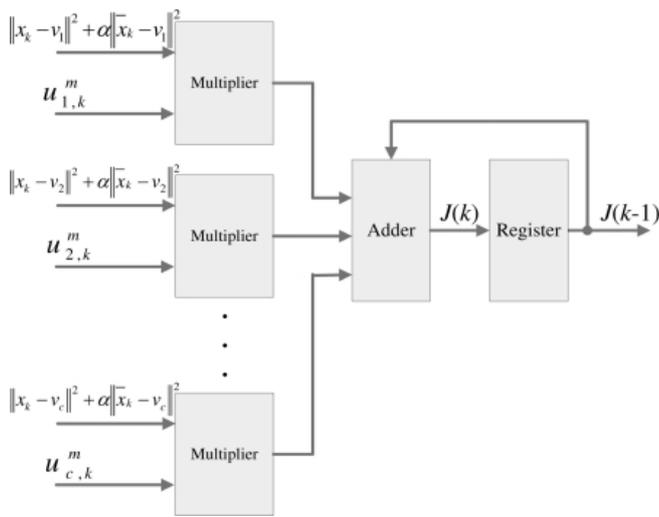
Fig. 6: The circuit for calculating cost function $J(k)$ for FCM-S.

We can view the membership coefficients, centroids and cost function for FCM-S as the extension of those for original FCM by replacing $\|x_k - v_j\|^2$ with $\|x_k - v_j\|^2 + \alpha\|\bar{x}_k - v_j\|^2$. Therefore, the membership coefficients updating unit, centroids updating unit and cost function computation unit for FCM-S also have similar architectures to those of their counterparts in original FCM. The circuits in FCM-S require only additional squared distance unit and adder for computing $\|x_k - v_j\|^2 + \alpha\|\bar{x}_k - v_j\|^2$.

3.6. The SOPC System Based on the Proposed Architecture

The proposed architecture is used as a custom user logic in a SOPC system consisting of softcore NIOS CPU, DMA controller and SDRAM, as depicted in Figure 13. The set of training vectors is stored in the SDRAM. The training vectors are then delivered to the proposed circuit by the DMA controller. The softcore NIOS CPU is running a simple software for FCM. It does not participate in the partitioning and centroid computation processes. The software only activates the DMA controller for the delivery of training vectors. The CPU then receives the overall distortion of clustering from the proposed circuit after the completion of DMA operation. The same DMA operation for delivering the training data to the proposed circuit will be repeated until the cost function J converges. The CPU then collects the centroid of each cluster from the proposed circuit as the clustering results.
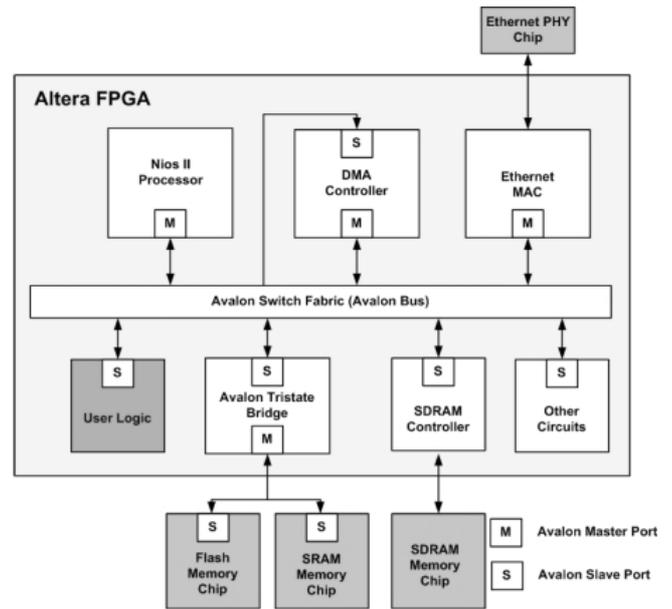
Fig. 7: The SOPC system for FCM-based image segmentation.

4.0   Experimental Results

This section presents some numerical results of the proposed FCM-S architecture for image segmentation. The design platform of our system is Altera Quartus II with SOPC Builder and NIOS II IDE. The target FPGA device for the hardware implementation is Altera DE2 board [20].

For sake of brevity, the images considered in this section are gray-level images. Each data point x_k represents a pixel with gray level values in the range between 0 and 255. For color images, each pixel x_k becomes a vector consisting of three color components: red, green and blue. In the proposed architecture, each data point x_k can be a scalar or a vector. Therefore, the proposed architecture can be directly applied to color image segmentation by implementing x_k as a 3-dimension vector.

The performance of the image segmentation is measured by segmentation error rate, which is equal to the number of misclassified pixels divided by the total number of pixels.

Section consists in the computing of the percentage corresponding to the approximations vectors and the details vectors. In the following table are shown the computed values of mean coefficients obtained using a fuzzy c mean for a malign and a benign case sample which taken from MIAS Database.

The original MIAS Database (digitized at 50 micron pixel edge) has been reduced to 200 micron pixel edge and clipped/padded so that every image is 1024×1024 pixels. This is very useful for medical research purpose. MIAS Database also provides character of background tissue, class of abnormality, Severity of abnormality and the coordinate and radius for it.

In the model, the resulting clusters to 7 are shown superimposed upon the mammogram data. As Fuzzy-C Means is an iterative algorithm, the centers of the clusters in new iteration are shown using different colors mimic the original image. Bold white is the final clustering and previous iterations are shown in gray spectrum edited mammogram. Although create different coefficients in superimposed cluster centers, it does a decent job of clustering this data. Besides the clustering created in different folder and saved as cluster 0 to 6 (for clustering = 7) in PNG format file. The result for breast cancer image will be shown in one of the crusted image. These improvements in the way we image and diagnose breast cancer. Furthermore, it helps the radiology and medical personal in their daily routine to analysis Breast cancer mammogram.

Practically speaking, no difficulties have ever been encountered, and numerical convergence is usually achieved in 10 - 25 iterations. This is depending on the grey distribution and the pixel size of the image. Whether local minima good clustering is another matter, for it is easy to obtain data sets upon which p_k minimizes globally with x_k visually unappealing substructure. To mitigate this difficulty, several types of cluster validity functional are usually calculated on each μ produced by FCM. Among the most popular are the partition by calculation of the centroid of the i-th cluster is achieved using equation (26). Result Figure 8 shown the iterations over segmentation result that computed with out training and supervision.
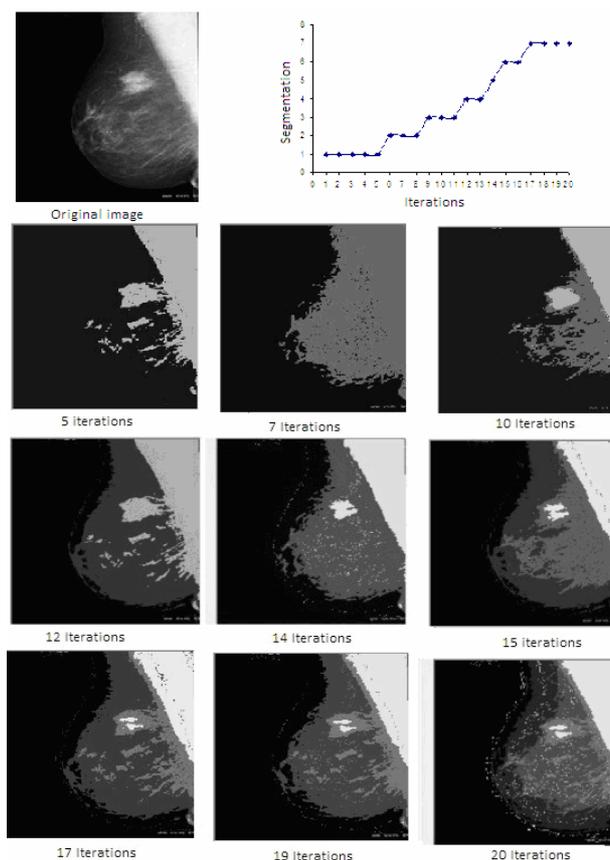
Fig. 8 Experiment mammograms image with different input Iterations and end segmentation.

We evaluated our algorithm on a series of images including the ones shown on the right side Figure 8. This continued by loading different breast cancer diagnosis form MIAS database. To see how our algorithm fares it terms of efficiency we conducted speed tests for all mammogram in MIAS Database by measuring how long they take until the final clustering (i.e. until the algorithms have converged). In order to arrive at figures which are independent of the hardware configured [19], we standardize them so that the fast algorithm is assigned by setting clustering 7 and iteration 25, produce high accuracy and higher precision for the root image created.
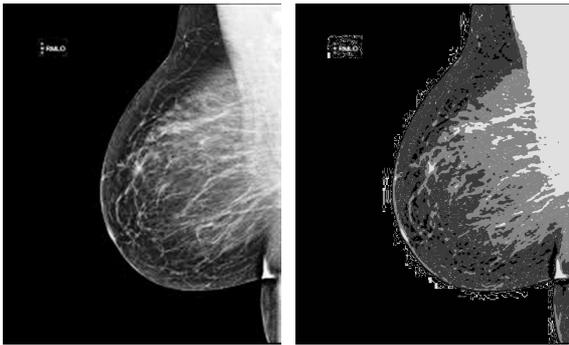
(a) Original Mammogram   (b) Segmented

(i) Cluster 7**(Target)**

Fig. 9 (a) is the original Mammogram image and 8 (b) is the Segmented image. The image is Cluster into five dominant regions as shown on 9 (c) -8 (i).

Result experiment for two different Severity of abnormality in breast cancer shown in graph in Figure 10 and Figure 11. The result for Benign abnormality is 98.4% while Malignant abnormality is 97.1%. Experiment conducted by compare the original diagnosis region which from MIAS Database with the image cluster (the breast cancer region).

(c) Cluster 1   (d) Cluster 2

(e) Cluster 3   (f) Cluster 4

Fig. 10 Result for Benign and Malignant breast cancer mammograms image segmented with different input Iterations.
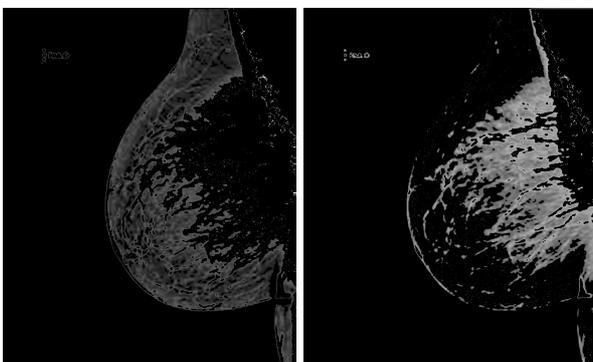
(g) Cluster 5   (h) Cluster 6
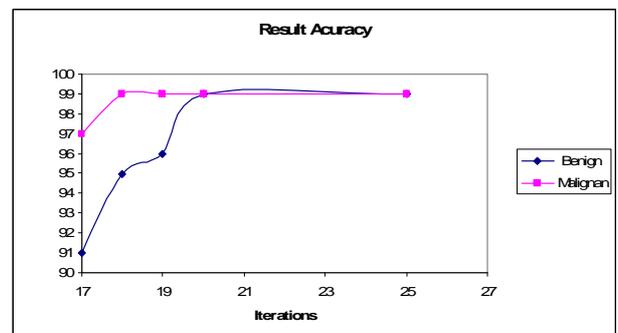
Fig. 11 Accuracy result for Benign and Malignant breast cancer.

5.0   Concluding Remarks

The proposed FCM-S architecture has been found to be effective for medical image segmentation. To lower the segmentation error rate, in the proposed architecture, the spatial information is used during the FCM training process. The architecture can also be designed for different values of degree of fuzziness to further improve the segmentation results. In addition, the architecture employs high throughput pipeline to enhance the computation speed. The n-th root circuits and inverse operation circuits in the architecture are designed by simple lookup tables and multipliers for lowering the hardware resource consumption. Early experimental results reveal that the proposed architecture is able to achieve segmentation error rate down to 5% for noisy images. Conduct for future analysis with the real Digital Mammogram Machine and this research predict much useful for Breast Cancer Prognosis. In addition, the SOPC architecture attains speedup up to 150 over its software counterpart. The proposed architecture therefore is an effective alternative for applications requiring real time image segmentation and analysis.

References:

[1] Lim SR, Halimah Y. Cancer Incidence in Peninsular Malaysia 2003-2005. The Third Report of the National Cancer Registry Malaysia. Kuala Lumpur: National Cancer Registry, Malaysia 2008.

[2] Bezdek, James C. Fuzzy sets in approximate reasoning and information systems. Kluwer Academic, 1999, ISBN: 0792385845.

[3] Ahmed MN, Yamany SM, Mohamed N, Farag AA, Moriarty T. A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data. IEEE Trans. Med. Imaging. 2002;21:193–199.

[4] Cai, Weiling, Songcan Chen, and Daoqiang Zhang. "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation." Pattern Recognition 40.3 (2007): 825-838.

[5] Keh-Shih Chuang, Hong-Long Tzeng, Sharon Chen, Jay Wu, Tzong-Jer. "Fuzzy c-means clustering with spatial information for image segmentation." computerized medical imaging and graphics 30.1 (2006): 9-15.

[6] Cannon R, Dave J, Bezdek J. Efficient implementation of the fuzzy c-means clustering algorithm. IEEE Trans. Pattern Anal. Mach. Intell. 1986;8:248–255.

[7] Cheng, Tai Wai, Dmitry B. Goldgof, and Lawrence O. Hall. "Fast fuzzy clustering." Fuzzy sets and systems 93.1 (1998): 49-56.

[8] Eschrich, S.; Jingwei Ke; Hall, L.O.; Goldgof, D.B.; , "Fast accurate fuzzy clustering through data reduction," Fuzzy Systems, IEEE Transactions on , vol.11, no.2, pp. 262- 270, 2003

[9] Kolen, John F., and Tim Hutcheson. "Reducing the time complexity of the fuzzy c-means algorithm." Fuzzy Systems, IEEE Transactions on 10.2 (2002): 263-267.

[10] Hwang, Wen-Jyi, Zhe-Cheng Fan, and Tsung-Mao Shen. "Unsupervised Image Segmentation Circuit Based on Fuzzy C-Means Clustering." CENICS 2012, The Fifth International Conference on Advances in Circuits, Electronics and Micro-electronics. 2012.

[11] Jesús Lázaro, Jagoba Arias, José L. Martín, Carlos Cuadrado, Armando Astarloa. "Implementation of a modified Fuzzy C-Means clustering algorithm for real-time applications." Microprocessors and Microsystems 29.8 (2005): 375-380.

[12] Li HY, Yang CT, Hwang WJ. Efficient vlsi architecture for fuzzy c-means clustering in reconfigurable hardware. Proceedings of The 4th International Conference on Frontier of Computer Science and Technology; Shanghai, China. 17–19 December 2009; pp. 168–174.

[13] Hauck S, Dehon A. Reconfigurable Computing: The Theory and Practice of FPGA-Based Computing. Morgan Kaufmann; San Fransisco, CA, USA: 2008.

[14] *NIOS II Processor Reference Handbook.* Altera Corporation; San Jose, CA, USA: 2011. Available online: http://www.altera.com/literature/lit-nio2.jsp (accessed on 20 December 2012)

[15] Lim SR, Halimah Y. Cancer Incidence in Peninsular Malaysia 2003-2005. The Third Report of the National Cancer Registry Malaysia. Kuala Lumpur: National Cancer Registry, Malaysia 2008.

[16] Bezdek, James C. Fuzzy sets in approximate reasoning and information systems. Kluwer Academic, 1999, ISBN: 0792385845.

[17] Ahmed MN, Yamany SM, Mohamed N, Farag AA, Moriarty T. A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data. IEEE Trans. Med. Imaging. 2002;21:193–199.

[18] Hwang, Wen-Jyi, Zhe-Cheng Fan, and Tsung-Mao Shen. "Unsupervised Image Segmentation Circuit Based on Fuzzy C-Means Clustering." CENICS 2012, The Fifth International Conference on Advances in Circuits, Electronics and Micro-electronics. 2012.

[19] Jesús Lázaro, Jagoba Arias, José L. Martín, Carlos Cuadrado, Armando Astarloa. "Implementation of a modified Fuzzy C-Means clustering algorithm for real-time applications." Microprocessors and Microsystems 29.8 (2005): 375-380.

[20] *Stratix II Device Handbook*; Altera Corporation: San Jose, CA, USA, 2011. Available online: http://www.altera.com/literature/lit-stx2.jsp (accessed on 20 December 2012).