IMPLEMENTATION OF MEDICAL IMAGE SEGMENTATION USING K-MEANS CLUSTERING TECHNIQUE ON FPGA

Smt.C.Chandrakala[1],T.S.Ghouse Basha[2]and K. Anjani[3]

[1]Assistant Professor, Department of Electronics&Communications Engineering, K.O.R.M. College,Kadapa,Andhrapradesh,India.

[2]Associate Professor&Head, Department of Electronics&CommunicationsEngineering,K.O.R.M.College,Kadapa,Andhrapradesh,India.

[3]Student, M.Tech [vlsi],Department of Electronics&CommunicationsEngineering,K.O.R.M.College,Kadapa,Andhrapradesh,India.

Abstract - K-Means is an clustering algorithm that is most essential functional to distinctive applications together with color clustering and image segmentation. The dimension of cluster numbers in embedded systems, hardware architecture of hierarchical K-Means (HK-Means) is planned to maintain a maximum cluster number of 1024. A hierarchical memory structure is incorporated to offer a highest bandwidth of 1280 bit/cycle to giving out elements. Features such as video segmentation and color quantization can be executersupport on the planned HK-Means hardware works and associated works. The earlier K-Means architectures cannot make happy the costs of together the computational time and the hardware area. Inembedded systems cost isall the time essential and the effort between the computational time and the hardware area becomes severe particularly as the cluster number enlarge. The large cluster number is a specific design challenge for K-Means hardware architectures. To correct the problem, a new hardware architecture support on hierarchical K-Means (HK-Means) is planned. The presentarchitecture containa hierarchical memory structure to accumulate the cluster centroids for distance calculations and binary-tree traversal are used to identify the distant centroid operations in pipeline.

Index term:

k-means,Imagesegmentation,Clustering,FPGA,

2. INTRODUCTION

Image segmentation method and a process which used to separate the image into different parts of region and take out the interested target. To clarify the level of the image segmentation in image processing, we have introduced methods, algorithms, tools, equipment of image segmentation into an overall framework. With the development of computer processing capacity and the improved application of color image, the image segmentation are more and more concerned.

Image segmentation is a key step from the image processing to image analysis, it occupy an essential place. It is the source of target appearance and has important consequence on the feature measurement. It needs to take out and divide them in order to recognize and examine the object, on this basis it will be possible to further use for the target.

Segmenting an image into similar parts is important for low level image understanding.

Many formulations of the segmentation task have been recommended over the years. While axiomatic functional, such as the Mumford-Shah functional are hard to execute and analyze, graph-based alternatives often require artificial measures on the problem. The latter are typically easy to optimize and execute at the expense of giving up some required properties.

ISSN: 2278 – 7798

Image segmentation is broadly used for object recognition and classification, scene understanding, action classification, and other visual information analysis tasks. The level set framework, when utilized per se, is geared towards two-region image segmentation. To require this limitation, different methods were developed and used them however,necessitate managing multiple level Set functions. Some connect a level set function with each image region, and develop these functions in a coupled manner.Others perform hierarchical segmentation, by iteratively splitting previously obtained regions using the conventional level set framework . These methods too require coupled level set evolution, so that the resulting regions do not develop gaps or overlaps. It is also possible to use a smaller number of level set functions, say n, and segment an image into 2n regions.

A different method to image segmentation consists to describe it as a different labelling problem and correct it using graph-cuts or convex relaxation algorithms. These processwere used for certain type of problems, but at present are not using for various energy measures, for certain, the elastic term in (integral of the squared curvature of the evolving contour). In addition, they frequently require knowing the number of regions a priori.

An important goal of medical image processing is to transform raw images into a numerically symbolic form for improved representation, evaluation, and/or content based search and mining. Annecessary step in this transformation is the segmentation of the target structures; that is, based on given homogeneity criteria, the task is the image separation into regions, which, in medical images, are frequently the target anatomic regions (foreground) and their surroundings (background). After this segmentation, the exact shape and appearance features of the targets can be calculated, and based on the application; they can be used for clinical evaluation, pattern analysis, and/or knowledge discovery.

A challenging problem is to segment regions with boundary insufficiencies, i.e., missing edges and/or lack of texture contrast between regions of interest (ROIs) and background. on a exact category of segmentation methods, namely the deformable models. The major cause why these move towards have been generally used in medical image computing is their robustness, mainly due to the models' constraints, as we explain throughout this chapter. We aim at giving the reader an instinctive but also mathematical description of these model-based methods, we describe their implementation aspects, and then we detail on superior methods that incorporate shape and appearance models for robust yet perfect medical image segmentation.

For instance, in ultra-sound images, various regions are determined by region contrast, in terms of the strength speckle density/distribution, The core model definition is independent from the features used in the external force terms frequently ; in most cases, the image features are application-related, i.e., their choice depends on the image modality. while the edge information, in its description as the image gradient, is frequently too poor to be used. On the other hand, magnetic resonance (MR) and computerized tomography (CT) images have enough gradients for edge information to be used in segmentation. Deformable models are classified into two general approaches, the parametric and the geometric models, depending on how the model is distinct in the shape domain.

3.Procedure for Paper Submission

A. Review stage

The segmentation algorithms ,are used to know important hardware features to achieve the high performance. It based on general-purpose microprocessors have to be of low computational complexity in order to reach processing speeds. A solution to execute composite algorithms is to put together a custom hardware. It may take months to expand and verify a design, and Application-Specific Integrated Circuits (ASICs) incurs costs ranging from hundreds to hundreds of thousands of dollars.

B. Final stage

The present architecture contain a hierarchical memory structure to store the cluster centroids for distance calculations and binary-tree traversal are used to compute the nearest centroid operations in pipeline. The expansion of reconfigurable logic hardware in the form of Field-Programmable Gate Arrays (FPGAs) agree the designs to be rapidly developed and prototyped at moderately low cost. The simple of design related with software and the presentation connected with hardware can be attain using FPGA.

C.Figures

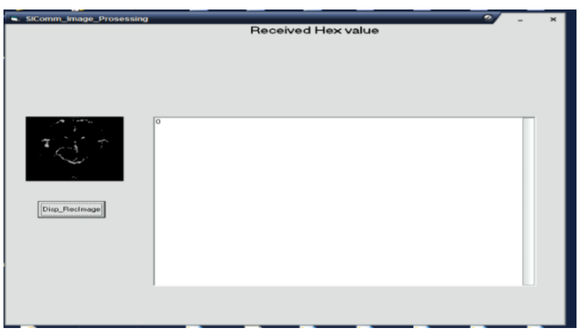
RESULTS:In Three parts experiments out comes contain the first part is the algorithm verification that

ISSN: 2278 – 7798

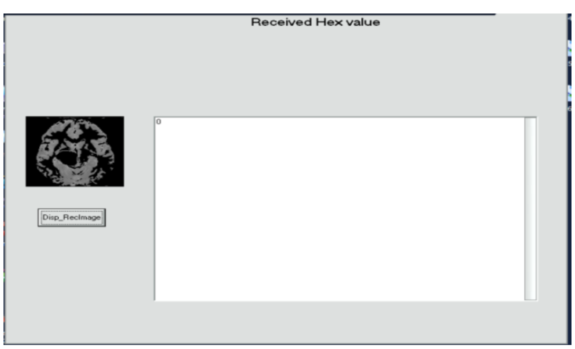
will chose to be the testing applications of the projected hardware. In color clustering, each and every input vector that is having three dimensions, which equals to the three color channels of a 24-bit full-color pixel. After HK-Means clustering, the color of every image pixel is correspond to the color of its corresponding centroid, and the development can also be treated as color quantization. The larger the cluster number K, the more image segments are create.

1.Input MRI Image

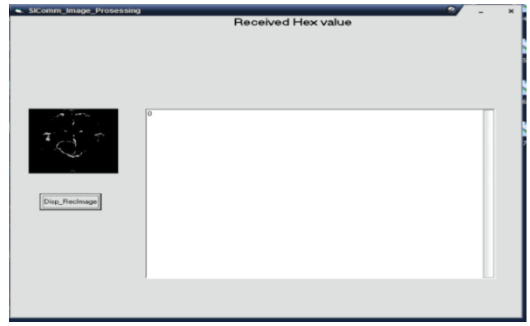
2.Input CT image



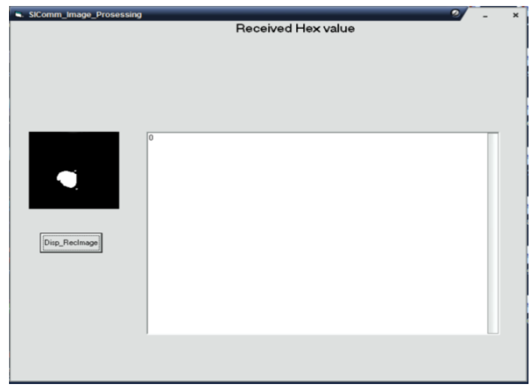
3.CT& MRI image



4.Edge Detection Image



5.SegmentedTumour Output



4.K-MEANSCLUSTERINGSEGMENTATION

K-means is one of the simplest unsupervised learning algorithms that resolve the well known clustering problem. The procedure follows a simple and easy way to classify a theen data set through a certain number of clusters (assume k clusters) fixed a priori.To define k cancroids of main idea , one for each cluster. These cancroids'supposed to be located in a cunning way because of different location causes different result. At this point we need to re-calculate k new centroids as barycenters of the clusters resulting from the earlier step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid.This algorithm aims at minimizing an objective function, in this case a squared error function. The objective function where is a selected distance measure between a data point and the cluster centre is anpointer of the distance of the *n* data points from their individual cluster centres.

K-means Algorithm

The algorithm uses a similarity metric to assign all documents to one of *k* clusters. The clusters are signifyas an normal of all documents restricted within the cluster. This average can be thought of as the centroid of the cluster.A simple two dimensional case for K-means clustering is shown The K-means algorithm set with $k = 4$ results in four clusters represented by *A*, *B*, *C*, and *D*.The K-means

ISSN: 2278 – 7798

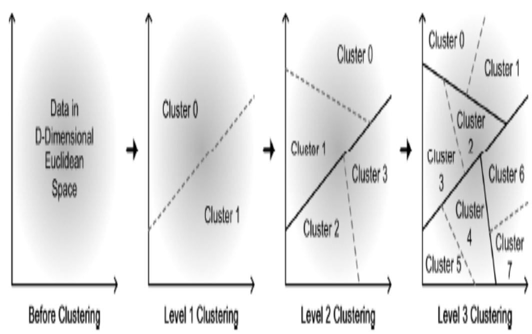
algorithm operates as follows:1. Using document vectors, *di 2 D*, to a cluster using an initial seed.

2. By using cluster centroids, *C*, from original document assignments.

3. For every document *d 2 D*(a) Recalculate distances from document *di* to centroids (*C*1,*C*2,...,*Ck*), and detectnearest centroid *Cmin*.

(b) Move document *d* from present cluster *Ck* into new cluster *Cmin*and re- calculate the centroid for *Ck*and *Cmin*.

FIG 6:of the HK-Means clustering process.



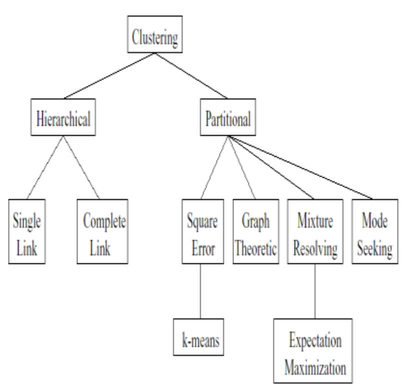


Fig 7:Levels Clustering Technique

5.HK-MEANS ALGORITHM

Hierarchical clustering method uses that create a tree structure or a dendrogram in the clustering process and there are top-down and bottom-up clustering move toward. Since there are no clustering methods that are appropriate for all the problems, many complementary, where HK-Means refers to a top-down and disturbing hierarchical clustering algorithm that adopts K-Means clustering with cluster number k=2 in each stage. The HK-Means concept is frequently functional to decrease the computational time in the software algorithm when the cluster number is large . it shows an illustration of the projected HK-Means clustering, which is a method that is appropriate for hardware design. The clusters are split into two recursively in the Euclidean space, and K-Means clustering with k=2is performed nearby in every level based on the clustering results of the previous level. An example of the binary-tree representation of clusters and 14 centroids are stored to handle the maximum cluster number k=8. It is assumed that the cluster number of HK-Means satisfies the condition that, and the steps in the HK-Means algorithm are stated as follows.

TABLE I
HIERARCHICAL MEMORY SPECIFICATIONS

Hierarchical Level	Memory Size (Bit × Word)	Activating Condition	Implementation Method
1	128 × 1	Always Activated	Register/Register File
2	128 × 2	$K > 2$	Register/Register File
3	128 × 4	$K > 4$	Register/Register File
4	128 × 8	$K > 8$	Register/Register File
5	128 × 16	$K > 16$	SRAM/Register File
6	128 × 32	$K > 32$	SRAM
7	128 × 64	$K > 64$	SRAM
8	128 × 128	$K > 128$	SRAM
9	128 × 256	$K > 256$	SRAM
10	128 × 512	$K > 512$	SRAM

.

IMPLEMENTATION OF CLUSTERING IN HARDWARE

Connected to thesystem bus, the “Data Memory” is used to store the feature data that are remove for multimedia applications and the feature data are regarded as input vectors for HK-Means clustering. Where the CPU, the external memory and other SIPs share the same bus resources with the HK-Means hardware.The“Data Memory” can offer input vectors to the HK-Means hardware with the throughput of 1 vector/cycle for iterative vector processing which put away a huge amount of bandwidth. The maximum vector dimension is set to 8, the bit length of each dimension D isset to 8, and the maximum cluster number K is set to 1024.The “Hierarchical Memory” is used to store a total of 2046 centroid vectors and offer a maximum of 10 centroid vectors to its neighboring modules at the same time.

ISSN: 2278 – 7798

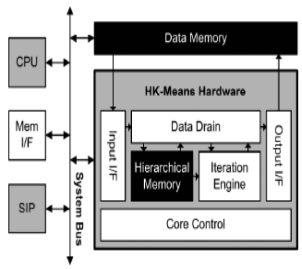


Fig 8: Data memory.

The "Data Drain" accept the input vectors from the "Data Memory" and include 10 sets of distance calculators to compute the nearest centroids of inputvectors in pipeline. The "Iteration Engine" works the iterative clustering process of K-Means with cluster number K=2 and communicates with the two modules mentioned above.Overview of the system environment and the planned HK-Means architecture, in which three main modules are included: the "HierarchicalMemory," the "Data Drain," and the "Iteration Engine."

MEMORY COST ANALYSIS

Two possible implementation methods of the memory architecture of HK-Means. The major one is the "Label-Based Memory," which means that the results of the nearest centroid computations of input vectors in each hierarchicallevel are accumulate in the memory label Based Memory" and the "Centroid-Based Memory" is shown in Fig. It is observed that the "Label-Based Memory" has lower memory cost than the "Centroid-Based Memory" when the data number is small, but the "Label-Based Memory" does not scale well when the data number is large. The "Centroid-Based Memory" outperforms the "Label- asedMemory." In large cluster number and preserve the flexibility for large data number, the "Centroid-Based Memory" is adopted in the proposed HK-Means hardware architecture.

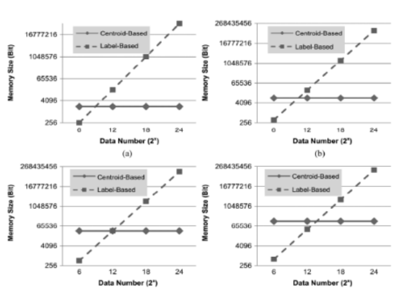


Fig 9:Memory size analysis of the "Centroid-Based Memory" and the "Label-Based Memory:" (a) luster number K=2^4=16 (b) Cluster number K=2^6=64 (c) Cluster number K=2^8=256. (d) Cluster number K=2^10=1024 the bit length of each dimension are set to D=8_and B=8 respectively.

DATA DRAIN

The "Data Drain" is a module that receives input vectors and computes the nearest centroids of them. Since there are 10 levels of centroids in the "Hierarchical Memory",a maximum of 10 nearest centroid operations can be execute to attain the results of input vectors. Because the computationsof the clustering with 1024 clusters are laborious, 10 sets of the "Traverse Processing Element (PE)" are applied to the "Data Drain" to accelerate the speed and the interface design of the"Traverse PE 10" is specialized to serve the functions in different states.

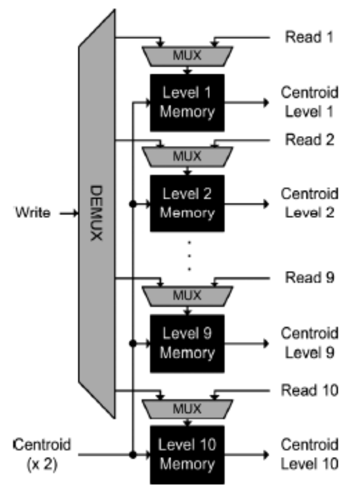


Fig.:10Architecture of the "Hierarchical Memory

Connected to different levels of the "Traverse PEs," the "Iteration Selector" and the "Output Selector" are stimulate in the INPUT state and the END state, respectively. In the INPUT state, the input vectors to the "Traverse PE 1 to 10" can be selected by the "Iteration Selector" according to the current level of HK-Means clustering. In the END state, the output vectorsfrom the "Traverse PE 1 to 10" can be selected by the "Output Selector" according to the cluster number K. In order to signify the cluster labels based on , the "Centroid Delay Line" is attached to the "Output Selector" to store the output centroid information from the "Traverse PE" in the earlier level.

ISSN: 2278 – 7798

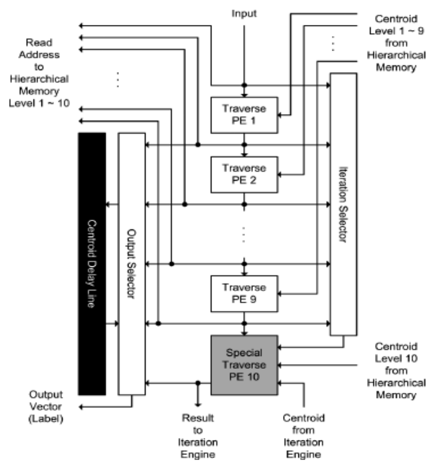


Fig.:11. Architecture of the "Data Drain."

6.IMPROVED K MEANS

Original K-means algorithm points as initial clustering centers, different points may attain different solutions. In order to reduce the sensitivity of original point assciatedtwo solutions generated by clustering sample drawn from the original dataset and itself using K-means respectively, the location of clustering centroids of these two are almost similar. So, the sample-based method is relevant to treat original circumstances [*I. In order to lessen the influence of sample on choosing original starting points following procedures are employed. First, drawing multiple sub-samples (say J) from original dataset (the size of each sub-sample is not more than the capacity of the memory and the sum for the size of J sub-samples is as close as possible to the size of original dataset) . Second, use K-means for each sub-sample and generate a group of mediods respectively. Finally, comparing J solutions and choosing one group having minimal value of square-error function as the developed initial points.

To avoid dividing one big cluster into two or more ones for adopting square-error criterion, we suppose the number of clustering is *K'* (K > K, K' depends on the balance of clustering quality and time). Near some extremism.Consequently, re-clustering the dataset through K-means with the chosen original conditions would produce K' mediods, then merging K' clusters (which are nearest clusters) until the number of clusters reduced to k.

PSEUDO CODE FOR IMPROVED K MEANS

Algorithm: Improved K-means(S, k), $S=\{x_1,x_2,\cdots,x_n\}$.

Input: The number of clusters $K'(K' > K)$ and a dataset containing n objects (Xi).

Output: A set of k clusters (Cj) that minimize the squared-error criterion

Begin

Multiple sub-samples **{SI, S2,** . . . ,Sj };

for m=l to j do

K-mean(Sm, K'); ***//executing k-means,***

I/ produce K' clusters and j groups.

Compute

$$J_c(m)=\sum_{j=1}^{k'}\sum_{X_i\in C_j}\left|X_i-Z_j\right|^2$$

choose min{J,} as the refined initial points $Z_j, j\in\left[1,K'\right]$;

K-means(S, K');

//executing k-means again with chosen initial

//producing K' mediods.

Repeat

Combining two near clusters into one cluster, and

recalculate the new center generated by two centers

merged.

Until the number of clusters reduces into k //Merging （K'→K）

End

7.CONCLUSION:

To improve the K-means algorithm is a solution to handle large scale data that can select initial clustering center with determination to decreasethe sensitivity to isolated point to avoid dissevering big cluster. By using this technique locating the original seed point is easy and which will give more correct and high-resolution result. By using different techniques we can study or associate the results and find out which technique gives higher resolution.

ISSN: 2278 – 7798

REFERENCES:

[1] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.

[2] L. D. Cohen, "On active contour models and balloons," *CVGIP, Image Understand.*, vol. 53, no. 2, pp. 211–218, Mar. 1991.

[3] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Gradient flows and geometric active contour models," in *Proc. Int.Conf. Comp. Vis.*, Jun. 1995, pp. 810–815.

[4] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jun. 1995, pp. 694–699.

C.Chandrakala is presently working as an Assistant Professor in the Department of Electronics and Communication Engineering in KORM College of Engineering, Kadapa. She carried out her MTech project work in jntuananthpur, and working in teaching field since 7 years in different cadres. She received BTech and MTech from the Department of Electronics and Communication Engineering from JNTUA.

T.S. GhouseBasha is presently working as an Associate Professor and HOD in the Department of Electronics and Communication Engineering in KORM College of Engineering, Kadapa. He carried out his MTech project work in Defence Research and Development Laboratory, Hyderabad and working in teaching field since eleven years in different cadres. He received his BTech and MTech from the Department of Electronics and Communication Engineering from JNTU University and Nagarjuna University respectively. He has submitted his Ph D thesis in microwave antennas to JNTUA. His areas of interest include microwave antennas, digital signal processing and mobile communications.

K.Anjani, Student, is currently pursuing her M.Tech VLSI., in ECE department from KORM Engineering College, Kadapa. She completed B.Tech in Electronics and Communication Engineering in KandulaLakshumma Memorial College of Engineering For Women. Her interest areas are VLSI systems,& wireless communication.

ISSN: 2278 – 7798