SPATIAL KERNEL-BASED GENERALIZED C-MEAN CLUSTERING FOR MEDICAL IMAGE SEGMENTATION

LEE SONG YEOW

UNIVERSITI SAINS MALAYSIA

2010

SPATIAL KERNEL-BASED GENERALIZED C-MEAN CLUSTERING FOR MEDICAL IMAGE SEGMENTATION

by

LEE SONG YEOW

**Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science**

December 2010

DECLARATION

Name: **LEE SONG YEOW**

Matric No: **PCOM 0041/09**

School: **SCHOOL OF COMPUTER SCIENCES, UNIVERSITI SAINS MALAYSIA**

Thesis Title:

Spatial Kernel-based Generalized C-mean Clustering for Medical Image Segmentation

I hereby declare that this thesis I have submitted to **SCHOOL OF COMPUTER SCIENCES** on **DECEMBER 2010** is my own work. I have stated all references used for the completion of my thesis.

I agree to prepare electronic copies of the said thesis to the external examiner or internal examiner for the determination of amount of words used or to check on plagiarism should a request be made.

I make this declaration with the believe that what is stated in this declaration is true and the thesis as forwarded is free from plagiarism as provided under Rule 6 of the Universities and University Colleges (Amendment) Act 2008, University Science Malaysia Rules (Student Discipline) 1999.

I conscientiously believe and agree that the University can take disciplinary actions against me under Rule 48 of the Act should my thesis be found to be the work or ideas of other persons.

Students Signature: Date: **February 17, 2011**

Acknowledgement of receipt by: Date: **February 17, 2011**

ACKNOWLEDGEMENTS

First of all, I would like to thank to Assoc. Prof Mandava Rajeswari for willing to take the trouble to be my supervisor, guiding me through out the process of preparation and writing of the thesis. She have given me a lot of encouragements, supports, and advice over my research.

I would like to express my appreciation to Dr. Dhanesh Ramachandram and Dr. Bong Chin Wei for giving me advice and help on my research. Nevertheless, Mr. Ong Kok Haur, Miss Lim Khai Yin, Mr. Bhavik Anil Chandra and Mr. Tan Ewe Hoe for willing to spend time listen,advice and discuss with me for my research.

Specially thanks to Miss Lee Liew Ean, my beloved for encouraging me all the way. Last but not least, to my family specially my father and mother, thank you for all your support when ever I need, without both of you, I wont make it so far.

TABLE OF CONTENTS

CHAPTER 6 – CONCLUSION

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

x

IPS   Institut Pengajian Siswazah

PPSK   Pusat Pengajian Sains Komputer

USM   Universiti Sains Malaysia

KM   K-means

RF   Radio Frequency

KHM   K harmonic-means

FCM   Fuzzy C-means

PCM   Possibilistic C-means

MRI   Magnetic Resonance Imaging

FGcM   Fuzzy Generalized C-Means

PGcM   Possibilistic Generalized C-Means

MIPAV Medical Image Processing, Analysis and Visualization

DICOM Digital Imaging and Communications in Medicine

CT   Computed Tomography

EPI   Extreme Physical Information

KFGcM Kernel-based Fuzzy Generalize C-means

KPGcM Kernel-based Possibilistic Generalize C-means

KGcM   Kernel-based Generalize C-means

SI   Similarity Index

TPF   True Positive Fraction

EF   Extra Fraction

MIPAV Medical Image Processing, Analysis and Visualization

MICCAI Medical Image Computing and Computer Assisted Intervention

WML   White Matter Lesion

FLAIR Fluid Attenuated Inversion Recovery

LIST OF SYMBOLS

lim   limit

θ angle in radians

PENGELOMPOKAN C-MIN TEITLAK BERASASKAN RUANG KERNEL UNTUK SEGMENTASI IMEJ PERUBATAN

ABSTRAK

Segmentasi imej merupakan salah satu tugas penting yang telah dibangunkan secara pesat sejak beberapa dekad yang lalu. Segmentasi imej dapat membantu dengan membahagi imej kepada kumpulan kecil dan menggabungkannya semula untuk membentuk hasil yang bermakna. Terdapat banyak algoritma yang telah dicadangkan untuk menyelesaikan masalah segmentasi, namun cabaran utama dalam pemprosesan imej perubatan adalah melakukan segmentasi dalam kehadiran ketidakhomogenan keamatan dan hingar.

Tesis ini dimotivasikan oleh masalah ini dan mencadangkan satu kaedah segmentasi terubahsuai yang menyatukan ciri baru, yang boleh meningkatkan kelasakan terhadap ketidakhomogenan keamatan dan hingar, di samping menonjolkan ketidaksamaan di antara kawasan-kawasan yang mempunyai variasi keamatan yang rendah. Ciri ini disatukan untuk memfomulasi algoritma pengelompokan c-min teritlak berasaskan ruang kernel yang baru. Algoritma ini menangani segmentasi kawasan-kawasan yang mempunyai ketidakhomogenan keamatan dan hingar.

Dua eksperimen telah dijalankan dalam tesis ini. Eksperimen pertama menggunakan imej-imej otak pengimejan magnetik resonans (MRI) untuk mensegmentasikan lesi jirim putih. Eksperimen dijalankan untuk membuktikan kemampuan algoritma pengelompokan c-min teritlak dan fungsi ruang kernel untuk mensegmentasi dalam kewujudan ketidakhomogenan keamatan. Eksperimen kedua menggunakan imej tomografi berkomputer untuk mensegmentasi

tumor dan kawasan hati. Eksperimen dijalankan untuk membuktikan ketepatan kaedah yang dicadangkan bagi mensegmentasi tumor hati.

Hasil eksperimen untuk kaedah yang dicadangkan dinilai dengan menggunakan matriks penilaian piawai, pertindihan kesatuan (Jaccard) dan pertindihan min (Dice). Hasil eksperimen menunjukkan bahawa kaedah yang dicadangkan dapat mensegmenkan kawasan yang dikehendaki dengan ketapatan yang tinggi berbanding dengan penandaan secara manual yang dilakukan oleh pakar. Kaedah yang dicadangkan memberi hasil yang menggalakkan dengan purata kadar pebaikan sebanyak 7.5%.

SPATIAL KERNEL-BASED GENERALIZED C-MEAN CLUSTERING FOR MEDICAL IMAGE SEGMENTATION

ABSTRACT

Image segmentation is one of the important tasks that has been rapidly develop in pass few decades. It helps by dividing an image into smaller groups and joining them together to form a meaningful result. There are many algorithms proposed to solve image segmentation problem, but the main challenge in medical image processing is to perform segmentation in the presence of intensity inhomogeneity and noise.

This thesis is motivated by this problem and proposes a modified segmentation method that incorporate new feature, which is able to increase the robustness to intensity inhomogeneity and noise, also enhance the dissimilarity between regions with low variations in intensity. This features is integrated to formulate a new kernel-based generalize c-mean algorithm. The algorithm addresses the segmentation of regions with intensity inhomogeneity and noise.

Two experiments is conducted in this thesis. First experiment uses brain Magnetic Resonance Imaging (MRI) images to segment the white matter lesion. The experiment conducted to prove the ability of generalize clustering algorithm and kernel function to segment in the presence of intensity inhomogeneity. The second experiment uses liver Computed Tomography (CT) images to segment tumor and liver region. The experiment conducted to prove the accuracy of proposed method on liver tumor segmentation.

The experimental results of the proposed method are evaluated using standard evaluation matrices, the Union overlapping (Jaccard) and Mean overlapping (Dice). The results shows that the proposed method able to segment the region of interest in close correspondence to the manual delineation by expert. The propose method produces an encouraging average improvement rate of 7.5%.

CHAPTER 1

INTRODUCTION

1.1 Introduction to Image Segmentation

Image segmentation is the process of dividing image into regions based on some notion of similarity. It is an important task in many different domains. For medical imaging, image segmentation can assist doctors and medical experts to locate and identify tumors. For pattern recognition, image segmentation is used to identify objects and characters. Image segmentation plays an important roles in image processing. It is a crucial preliminary low-level image processing that is required to facilitate higher level image analysis such as description, recognition, quantitative measurement and visualization of region of interest. In medical imaging for example, image segmentation techniques assist medical experts in medical image analysis, such as tumor detection, study of anatomical structure, and treatment planning.

1.2 Segmentation by Clustering

There are several approaches to image segmentation. Similar to segmentation, clustering is a process where data is grouped or clustered based on some notion of similarity. Thus, it is not surprising that most researched clustering methods are primarily focused on image segmentation. Among the clustering methods, K-means (KM) clustering is one of center-based clustering techniques (MacQueen, 1967). KM clusters the data according to the distance of each data to the respective centers. It partitions data into k sets or clusters. At first, the initial cluster centres are randomly set. It is followed by measuring the distance between center and pixels, and assigns pixels to the center of the nearest center. Euclidean distance measurement

1

is used for computing distance between the data and cluster center. Then center of each cluster is readjusted. KM will go through a series of iteration of updating the centroid until a stage where changes of centroid reaches its minimum value. However, K-means has its limitation due to this simplicity. First of all, the clustering performance of KM is highly dependant on its initial center (Li et al., 2007). Each data point is given the same weight, as the iteration goes on, the centroid of the cluster tends to converge at a higher dense of intensity area thus resulting in local minimum. Secondly, KM employs Euclidean distance measurement which is very sensitive to noise. It treats noise as an individual pixels and will perform distance calculation between centers, and assign it to the nearest center.

With the introduction of fuzzy theory by Dunn (1973), a pixel can have the probability of belonging to other clusters. The fuzzy theory is extended by James (1981). He introduced the fuzzy objective functions. As a result, Fuzzy C-means (FCM) was introduced. FCM has been widely used for image segmentation especially in medical imaging. FCM is able to improve clustering results by reducing the uncertainty of pixels belonging to one class (Hassanien et al., 2009). Fig.1.1 shows the evaluation of fuzzy theory. Data point X_1 lies between 2 clusters. Distance between data point X_1 to cluster center C_1 is nearer compared to the distance between data point X_1 to cluster center C_2. Therefore the probability of data point X_1 belonging to cluster C_1 is greater than the probability of data point X_1 belonging to cluster C_2.

Another improvement to FCM is the introduction of Possibilistic C-means (PCM). PCM was introduced by Krishnapuram and Keller (1993), where its algorithm is robust to noise. PCM also helps in a clustering situation where the number of cluster is either known or unknown. PCM algorithm allows all data points to act independently to each other. PCM adapts well to cluster noisy images and also images with unknown number of clusters.

Among other improvements done by Menard et al. (2004), was the proposal of Fuzzy

$$P(x_1, C_1) = 0.7$$
$$P(x_1, C_2) = 0.3$$

$$x_1$$

$$C_1 \qquad C_2$$
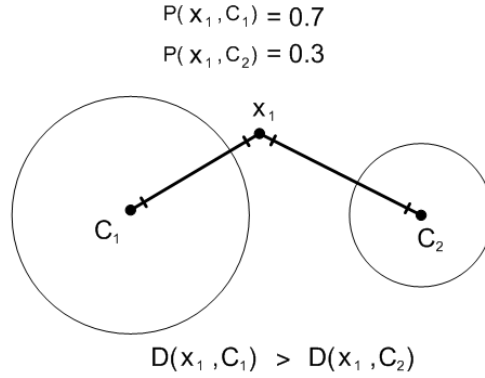
$$D(x_1, C_1) \; > \; D(x_1, C_2)$$

Figure 1.1: Data point X_1 has a distance of $D(X_1, C_1)$ to cluster center C_1 and distance of $D(X_1, C_2)$ to cluster center C_2, resulting $D(X_1, C_1) > D(X_1, C_2)$, therefore the probability of X_1 to cluster center $C_1, P(X_1, C_1)$ is higher than the probability of X_1 to cluster center $C_2, P(X_1, C_2)$

Generalized C-Means (FGcM) and Possibilistic Generalized C-Means (PGcM). These methods compensate for intensity inhomogeneity during the clustering process itself. Both algorithms employ Tsallis entropy (Tsallis, 2002), where the entropy is a non-extensive entropy. Tsallis entropy also includes long range interaction. Long range interaction means the influence of one data from a range of distance to another data.

1.3 Motivation

This research was particulary focused on MRI. Segmentation of MRI images is a challenging task owing to two main issues. The first issue is the intensity homogeneity. In MRI, the acquired images often appear inhomogeneous in intensity due to the non-uniformities in Radio Frequency (RF) field during the acquisition process. It results in a shading effect, a smooth variant of intensity across image (Franjo et al., 2007; Balafar et al., 2010). Intensities resolution between different anatomical structures are usually very small. This makes it extremely difficult to segment the anatomical structures based on the differences in intensity values.

In addition, MRI images may also be corrupted by noise. Noise can be mistakenly treated as a data point which has a sudden change in value within the neighborhood. Pixels on an

image is highly correlated, thus noise can be detected by considering the neighboring pixel relationship (Chuang et al., 2006). Fig.1.2 shows the pixels' values where within the group of pixels, the pixels have the similar range of values except one pixel that shows a sudden change of value. The sudden change of value on this pixel only occurs once within a small neighborhood. Therefore this pixel is considered as a noise to this set of pixels.

Figure 1.2: A pixel that has very big difference in intensity value compared to the surrounding pixels is considered as noise

Traditionally, MRI images are preprocessed to remove intensity inhomogeneities prior to segmentation. This research adapts the FGcM and PGcM proposed by Roullier et al. (2007) to combine the inhomogeneity correction and clustering into one step. In addition to this, this research hypothesized that by introducing kernel distance into clustering algorithms may overcome the problem of intensity resolution discussed earlier. This the proposed method would increases the interclass distance and decreases the intra-class distance in the clustering process. The objectives of this thesis are presented in the following section.

1.4 Objectives

- To propose a clustering algorithm that is robust to intensity inhomogeneities in MRI images.

- To incorporate kernel distance into the clustering algorithm to improve inter class separation.

1.5 Importance of Research

This research gives a new knowledge into image clustering area. A new modified clustering algorithm is proposed to segment MRI images. The proposed method is robust in the presence of noise and inhomogeneity.

1.6 Scope of Research

This work is limited to MRI images. Only FCM based clustering algorithm and Gaussian kernel distance measurement shall be studied in detail and verified.

1.7 Contribution

This research proposes a new clustering algorithm that is effective in segmenting MRI images and robust to noise and intensity inhomogeneity

1.8 Overview of Proposed Method

This research contains a few phases as shown in Fig.1.3. The procedure starts with identifying types of kernel function to be used in the research. It is followed by incorporating the kernel function into FGcM and PGcM clustering algorithm. Then the output of the algorithm is

evaluated using standard evaluation matrices. The process of clustering algorithm is carried out in stages. Image is input to the kernel-based FGcM and its output is used as the input for kernel-based PGcM. The output of the kernel-based PGcM will be evaluated using standard evaluation matrices.
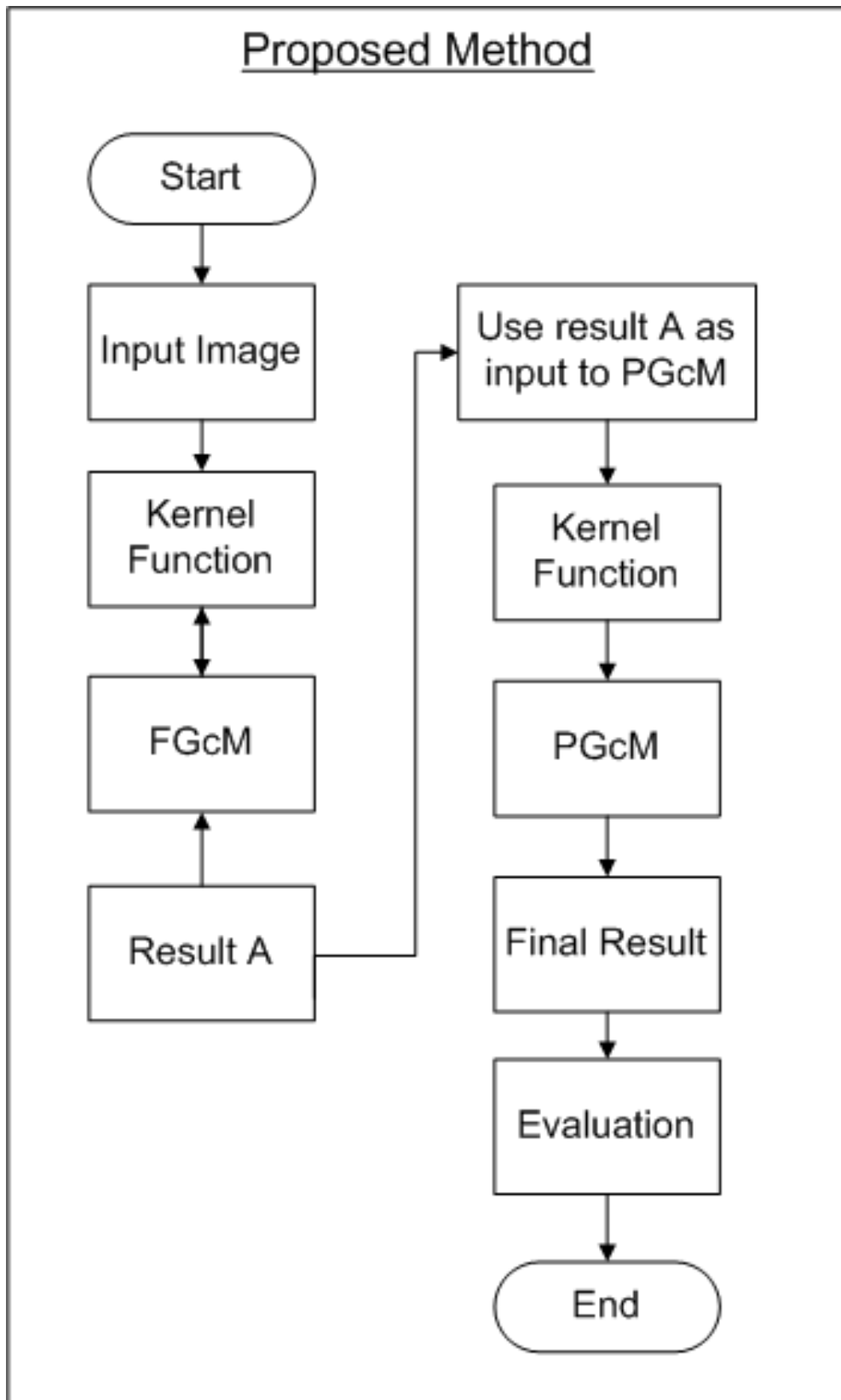
Figure 1.3: Overview of Proposed Method

1.9 Thesis Organization

The thesis is organized as follows:

In Chapter 1, the thesis will briefly introduce image segmentation and some background of image segmentation by using clustering approach. This chapter also discusses the motivation, scope, contribution and also the overview of the research.

In Chapter 2, the theoretical background of various types of fuzzy clustering methods that will be used in the thesis will be discussed. The chapter will start with the discussion of FCM algorithm, followed by PCM, FGcM and PGcM.Kernel function and spatial information will be covered later on. This chapter shows most of the formulation for different algorithms. In this chapter, most of the formulation of the aforementioned algorithms will be discussed.

In Chapter 3, the thesis will be explore the related literature. An introduction to image segmentation is given, followed by introduction of Fuzzy C-Mean (FCM) clustering algorithm. Next, Probabilistic C-Mean (PCM) is presented followed by the comparison of both FCM and PCM algorithms. Next section will introduce FGcM, PGcM and kernel-based distance measurement.

In Chapter 4, the thesis will describe the proposed work. The chapter starts with an explanation of the procedure to identify implement kernel function into FGcM and PGcM in detail. It is then followed by parameter setting for the segmentation process. The standard evaluation matrices is identified and the procedure of the evaluation method is to be explained later in the chapter.

In Chapter 5, the thesis will describes of the procedure of experiments to test the performance of proposed algorithm. There will be two experiments to be carried out. The detail of

the experiment is describe in this chapter. It is followed by showing the results based on the experiments carried out. Then, data analysis based on the results obtained. Discussion of the results and analysis is carried out later in the chapter.

In Chapter 6, the thesis will discuss the conclusion of the overall thesis and future work.

CHAPTER 2

THEORETICAL BACKGROUND

In this chapter, the theoretical background of the major concepts introduced in the research is discussed.

2.1 Introduction to FCM

This section explains the theory of FCM and the formulations involved in FCM algorithm.

Medical image intensities have a small dynamic range. This poses a problem during segmentation. In this scenario, each pixel may belong to more than one cluster with a fuzzy degree of membership that is widely accepted. Classical FCM was proposed by Dunn (1973) and the algorithm was further improved by James (1981) with fuzzy objective function. The FCM algorithm can be minimized by the following objective function:

$$J(U,V) = \sum_{k=1}^{N} \sum_{i=1}^{C} \mu_{ik}^{m} \|x_k - c_i\|^2 \tag{2.1}$$

where μ_{ik} represents the fuzzy membership of sample x_k in i^{th} cluster center where $\sum_{i=1}^{c} \mu_{ik} = 1$. $X = \{x_1, x_2, x_3 \ldots x_k\}$, where $k =$ image resolution of $M \times N$. $C = \{c_1, c_2, c_3 \ldots c_i\}$, and C is the number of cluster. Parameter m controls the fuzziness of the resulting partition or the extent of membership sharing between each other. In reduce the complexity of the formula, assume that $D_{ik}^2 = \|x_k - c_i\|^2$, where D_{ik} would be the distance between the x_k and c_i. This formula is

assigned into equation Eq.2.1 to replace $\|x_k - c_i\|^2$.

$$J(U,V) = \sum_{k=1}^{N} \sum_{i=1}^{C} \mu_{ik}^m D_{ik}^2 \tag{2.2}$$

from Eq.2.2, the membership and centroid function will be updated by using the following

formula:

$$\mu_{ik} = \frac{D_{ik}^{-2/(m-1)}}{\sum_{j=1}^{C} D_{ij}^{-2/(m-1)}} \tag{2.3}$$

where $1 \le k \le N$, $1 \le j \le C$ and

$$v_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m x_k}{\sum_{k=1}^{N} \mu_{ik}^m} \tag{2.4}$$

where $1 < i < C$

However, FCM suffers from singularity problems. Singularity occurs when one or more of

the distance $D_{ik} = 0$ at any iterate (Pal et al., 2005). Constraints to Eq.2.3, in assigning $\mu_{ik} = 0$

where distance $D_{ik} = 0$, resulting the centroid update, Eq.2.4 will ignore the existence of the

pixel. Krishnapuram and Keller (1993) proposed a solution to solve the problems.

2.2 Introduction of PCM

This section explains the theory of PCM and the formulations involved in PCM algorithm.

PCM was proposed by Krishnapuram and Keller (1993) in order to overcome the problems

in FCM, by relaxing the constraint $\sum_{i=1}^{C} \mu_{ik} = 1$ for the fuzzy membership, PCM allows the

membership of pixel, μ_{ik} to have a better representation to the ith cluster. Krishnapuram and

Keller (1993) proposed PCM model is formed by using following the objective function:

$$J(U,V) = \sum_{k=1}^{N} \sum_{i=1}^{C} \mu_{ik}^m D_{ik}^2 + \sum_{i=1}^{C} \gamma_i \sum_{k=1}^{N} (1 - \mu_{ik})^m \tag{2.5}$$

11

From Eq.2.5, the membership and centroid function will be updated using the following formula:

$$\mu_{ik} = \frac{1}{1 + \left(\frac{D_{ik}^2}{\gamma_i}\right)^{\frac{1}{m-1}}} \tag{2.6}$$

where $1 \leq k \leq N$, $1 \leq j \leq C$, γ_i is a suitable positive value that will discussed later.

$$v_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m x_k}{\sum_{k=1}^{N} \mu_{ik}^m} \tag{2.7}$$

where $1 < i < C$

The objective function in Eq.2.5 is minimized. PCM assigns higher membership values to for the nearest cluster center and lower membership to other cluster centers. The first term of Eq.2.5 demands that the distance of x_k to the closest cluster center, c_i should be as low as possible. The second term forces the membership to be as large as possible. The value of γ_i determines the distance at which the membership value of a point in a cluster becomes 0.5. It needs to be chosen depending on the range of the membership distribution for each cluster. γ_i value can be a fixed number with the constraint that all clusters are expected to be similar. Otherwise, γ_i relate to the overall size and shape of the cluster c_i. γ_i value can be obtained by following equation:

$$\gamma_i = 1\frac{\sum_{k=1}^{N} \mu_{ik}^m D_{ik}^2}{\sum_{k=1}^{N} \mu_{ik}^m} \tag{2.8}$$

This makes γ_i proportional to the average fuzzy intra-cluster distance of cluster c_i.

PCM overcomes the need to specify the number of clusters which means it can be initialized to a bigger number of cluster c_i^*, where $c_i \geq c_i^*$. This is very useful when the initialization of each cluster is unkown, in other words, PCM improve the over segmentation generated by FCM. PCM is robust in noise environment. This is especially important when it comes to medical imaging where noise is one of the major concerns. Unlike FCM which is sensitive to

noise, PCM allows the pixels or data points to behave nearly independently in the data set.

However, PCM is sensitive to the initialization of cluster center v_i. The generated output is greatly influenced by the value of initialization. The freedom to act independently in the data set results in clusters overlapping and producing weak clusters.

2.3 Introduction of FGcM

This section explains the theory of FGcM and the formulation involved for FGcM algorithm

Intensity inhomogeneity in image segmentation causes the pixels in the same cluster to have broad range of intensity values. This problem is addressed by introducing Tsallis entropy (Tsallis, 2002) which takes care of long and short range interaction. Long range interaction in Tsallis entropy considers a point far from prototype and increase centroid separation in the objective function. FGcM is modeled by following objective function:

$$J(U,V;Y) = \sum_{k=1}^{N} \sum_{i=1}^{C} \mu_{ik}^m D_{ik}^2 + \frac{1}{\lambda(m-1)} \sum_{k=1}^{N} \sum_{i=1}^{C} \mu_{ik}^m - \frac{1}{\lambda} \sum_{k=1}^{N} \gamma_k (\sum_{i=1}^{C} \mu_{ik} - 1) \qquad (2.9)$$

The first term of the objective function is the mean square term as given by Bezdek objective function (James, 1981). The second term defines Tsallis entropy for $\sum_{i=1}^{C} \mu_{ik} = 1$. The Eq.2.9 is converged to the original FCM for m tending to 1 when the Tsallis entropy converges to Shannon entropy. Tsallis entropy has been used to replace Shannon entropy to embrace the full range of possibilities from a non-extensive setting (Menard et al., 2003). The parameter m in Tsallis entropy is the measure of the degree of non-extensibility. Usually entropy is considered extensive, which means its values are proportional to the system size. For image segmentation, image resolution defines the system size. Tsallis statistics are based on generalized entropy. Its application deals with systems which have points with long range interactions when

entropy becomes non-extensive. For short range interactions, it is interchangeable with the Gibs-Boltzman statistic and converges to Shannon entropy. The third term in the Eq.2.9 is the probabilistic constraint, where it can be derived using the Extreme Physical Information (EPI) principle (Menard and Eboueya, 2002). It justifies the imposed constraint on the FCM method and provides a method to derive FCM objective function along with the probabilistic constraint. The minimization of the FGcM objection function can be done by minimizing the membership function:

$$\mu_{ik} = \frac{[1 + \lambda(m-1)D_{ik}^2]^{(\frac{-1}{m-1})}}{\sum_{i=1}^{C}[1 + \lambda(m-1)D_{ik}^2]^{(\frac{-1}{m-1})}} \qquad (2.10)$$

The cluster center of each clusters are updated with following centroid function:

$$v_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m x_k}{\sum_{k=1}^{N} \mu_{ik}^m} \qquad (2.11)$$

μ_{ik} represent the fuzzy membership of sample x_k in i^{th} cluster center where $\sum_{i=1}^{c} \mu_{ik} = 1$. $x = \{x_1, x_2, x_3 \ldots x_k\}$, where $k =$ image resolution. $V = \{v_1, v_2, v_3 \ldots v_i\}$ is a vector of ith cluster center for $1 < i < C$. $\lambda > 0$ plays the role of Lagrange multiplier.

2.4 Introduction of PGcM

This section explains the theory of PGcM and the formulation involved in PGcM algorithm

The same idea from the previous method 2.3 is extended to PGcM by Menard and Eboueya (2002). For images with low dynamic range coupled with intensity inhomogeneity, PGcM has a better performance than PCM where PGcM employs Tsallis entropy. PGcM is modeled by following objective function:

$$J(U,V;Y) = \sum_{k=1}^{N}\sum_{i=1}^{C}\mu_{ik}^m D_{ik}^2 + \frac{1}{\lambda(m-1)}\sum_{k=1}^{N}\sum_{i=1}^{C}[\mu_{ik}^m - \mu_{ik}] - \frac{1}{\lambda}\sum_{k=1}^{N}(\sum_{i=1}^{C}\mu_{ik}) \qquad (2.12)$$

The first term of the objective function is the mean square term as given by Bezdek objective function (James, 1981). The second term defines Tsallis entropy for $\sum_{i=1}^{C} \mu_{ik} = 1$. The objective function can be minimized by minimizing the membership function as follows:

$$\mu_{ik} = \frac{1}{[1 + \lambda(m-1)D_{ik}^2]^{(\frac{-1}{m-1})}} \tag{2.13}$$

μ_{ik} represent the fuzzy membership of sample x_k in i^{th} cluster center. $\lambda > 0$ plays the role of Lagrange multiplier. In PGcM, Pal et al. (2005) suggest to compute λ by using :

$$\lambda = \frac{\sum_{k=1}^{n} \mu_{ik}^m D_{ik}}{\sum_{k=1}^{n} \mu_{ik}^m} \tag{2.14}$$

2.5 The Kernel Function

This section explains the concept of kernel function and the formulation involved for FGcM algorithm.

From the literature Muller et al. (2001); Songcan and Daoqiang (2004); Kim et al. (2005); Li et al. (2007); Graves and Pedrycz (2010) , kernel function was used to transform pixel's intensity values from one feature space A to feature space B where feature space B is the separation of pixel's intensity values. Fig.2.1 shows the transformation in graphical form. X_k represents the pixel of x at the instance of k being transformed to feature space of ϕx_k. c_i represents the centroid of class i being transformed to another feature space. The kernel function transformed all pixels and centroid to another feature space.

With the constrains to Euclidean distance $D_{ik}^2 = \|x_k - c_i\|^2$, new distance calculation is re-formulated as $D_{ik}^2 = \|\phi(x_k) - \phi(c_i)\|^2$. The squared distance is computed in the kernel space
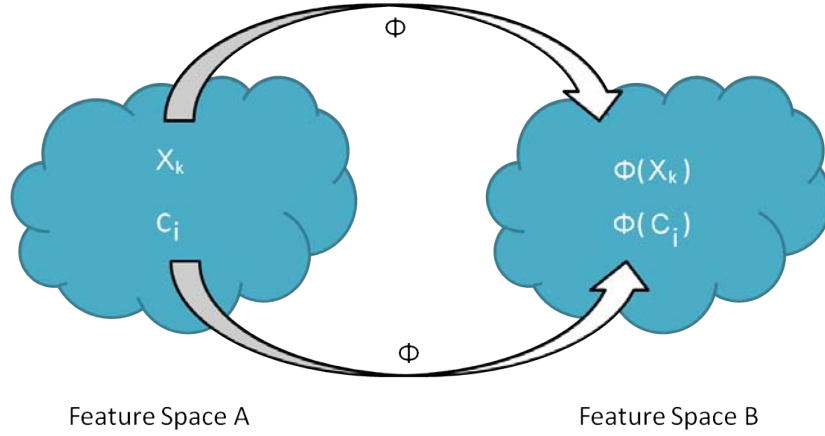
Figure 2.1: Feature space and kernel space

using kernel function where :

$$\|\phi(x_k) - \phi(c_i)\|^2 = [\phi(x_k) - \phi(c_i)] \cdot [\phi(x_k) - \phi(c_i)]$$

$$= \phi(x_k) \cdot \phi(x_k) + \phi(c_i) \cdot \phi(c_i) - 2(\phi(x_k) \cdot \phi(c_i)) \quad (2.15)$$

$$= [K(x_k, x_k)] + [K(c_i, c_i)] - 2[K(x_k, c_i)]$$

There are several kernel functions available:

$$K(x_i, x_j) = (x_i \cdot x_j + c)^d \quad (2.16)$$

where, $c \geq 0, d \varepsilon N$

$$K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2}) \quad (2.17)$$

where, $\sigma > 0$,

$$K(x_i, x_j) = \exp(-\frac{\sum \|x_i^a - x_j^a\|^b}{\sigma^2}) \quad (2.18)$$

where, $0 < b \leq 2$

$$K(x_i, x_j) = 1 - \tanh(-\frac{\|x_i - x_j\|^2}{\sigma^2}) \quad (2.19)$$

As Gaussian kernel which is used almost exclusively in literature Graves and Pedrycz

(2010); Muller et al. (2001), this research will use Gaussian kernel as the replacement of Euclidean distance. Then $K(x_k, x_k) = 1$, and Eq.2.15 can be simplify to:

$$[K(x_k, x_k)] + [K(c_i, c_i)] - 2[K(x_k, c_i)] = 2 - 2[K(x_k, c_i)]$$

$$= 2[1 - K(x_k, c_i)] \tag{2.20}$$

2.6 Spatial relationship

Ahmed et al. (2002) modified FCM objective function in such a way that a pixel can be assigned to the cluster if it is in the homogenous regions. In other words, Ahmed et al. (2002) proposed a modified objective function to increase the robustness of FCM. The objective function is defined as follows:

$$J(U, V) = \sum_{i=1}^{C} \sum_{k=1}^{N} \mu_{ik}^m \|x_k - v_i\|^2 + \frac{\alpha}{N_R} \sum_{i=1}^{C} \sum_{k=1}^{N} \mu_{ik}^m \sum_{x_r \varepsilon N_k} \|x_r - v_i\|^2 \tag{2.21}$$

where N_k stands for the set of neighbors that exist in a window around x_k and N_R is the cardinality of N_k. The parameter α in the second term controls the effect of the neighbors to the pixel.The objective function formulates the spatial constraint with the aim of keeping the continuity on the neighboring pixel values around x_k. The objective function $J(U, V)$ can be minimized by minimizing the membership μ_{ik}. The membership function is shown as follows:

$$\mu_{ik} = \frac{(\|x_k - v_i\|^2 + \frac{\alpha}{N_R} \sum_{r \varepsilon N_k} \|x_r - v_i\|^2)^{\frac{-1}{(m-1)}}}{\sum_{j=1}^{C} (\|x_k - v_i\|^2 + \frac{\alpha}{N_R} \sum_{x_r \varepsilon N_k} \|x_r - v_i\|^2)^{\frac{-1}{(m-1)}}} \tag{2.22}$$

where $1 \le k \le N$, $1 \le j \le C$ and

$$v_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m (x_k + \frac{\alpha}{N_R} \sum_{x_r \varepsilon N_k} x_r)}{(1 + \alpha) \sum_{k=1}^{N} \mu_{ik}^m} \tag{2.23}$$

where $1 < i < C$. The shortcoming of the Eq.2.22 and Eq.2.23, because of the complexity of the equation, is that it takes more computational time for the neighborhood terms as compared to FCM. Another technique to incorporate the spatial information and also to reduce the computational time is by calculating the probability of neighboring pixels to its pixel is proposed by Chuang et al. (2006). They utilized the standard FCM algorithm by exploiting the spatial information with the following spatial function:

$$h_{ik} = \sum_{j\varepsilon NB(x_k)} \mu_{ij} \tag{2.24}$$

where $NB(x_k)$ represents the square window center on the pixel x_k in the spatial domain.
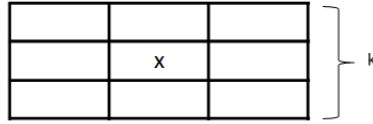
Figure 2.2: Square windows $NB(x_k)$ on pixel x_k

value of h_{ik} is large if majority of the neighboring pixels belongs to the same cluster. The spatial function is then incorporated with the membership function to be the new membership function:

$$\mu_{ik}^* = \frac{\mu_{ik}^p h_{ik}^q}{\sum_{j=1}^{C} \mu_{ij}^p h_{ij}^q} \tag{2.25}$$

where p and q are parameters to control the relative importance of both function. Cluster center will be updated using Eq.2.4 by replacing the μ_{ik} to μ_{ik}^*.

The extension of Eq.2.24 is done based on modifying the membership function(Mandava et al., 2010). They exploit the spatial function Eq.2.24 to a new spatial function as shown:

$$h_{ik} = \frac{\sum_{t \varepsilon \Omega_k} \mu_{it} \beta_t}{\sum_{j=1}^{C} \sum_{t \varepsilon \Omega_k} \mu_{it} \beta_t} \tag{2.26}$$

Where Ω_k represents a square window centered on pixel x_k in the spatial domain. h_{ik} having

the same function as the representation of the probability of pixel x_k belonging to ith cluster.

Spatial information is the largest if all of its neighboring pixels belong to ith cluster. β_t is

the contribution factor of the neighbor x_t to the overall spatial dissimilarity. The nearer the

neighbor is, the stronger the interaction becomes, and the bigger the contribution will be. β_t is

a non-increasing function of the distance between sites j and t and it is defined as follows:

$$\beta_t = \frac{1}{1 + \exp(\theta \cdot \|j - t\|)} \tag{2.27}$$

Where the coefficient θ determines how fast the interaction between two sites diminish with

the increasing of the distance between them. The Eq.2.26 is then used to update Eq.2.25 where

the results are shown to be better.

CHAPTER 3

LITERATURE REVIEW

In this chapter, the thesis will focus on various types of fuzzy clustering algorithms. This chapter will explore FCM and previous work that has been done on it for improvement. Next, the introduction of PCM as an alternative clustering method will be included. After that, the chapter will introduce the FGcM and PGcM as a new approach to clustering method on image segmentation. Finally, Tsallis entropy involving long range interaction will be discussed in detail followed by the kernel function.

3.1 Literature Review

This section focuses on FCM and some of its variations. FCM was originally proposed by Dunn (1973). With the introduction of fuzzy concept, a pixel can have the probability of belonging to more than one cluster. The FCM algorithm is further improved by James (1981) upon the introduction of fuzzy objective function algorithm. FCM clustering algorithm is a popular clustering model which has been widely used in image segmentation. It is an iterative algorithm which assigns pixels to each category by using fuzzy membership to produce the optimal cluster. FCM maximizes the inter-cluster distance while at the same time minimizing intra-cluster distance. FCM clustering algorithm is minimized by an objective function in Eq.2.2. The membership for pixels to cluster center is calculated using Eq.2.3. The centroid values are iteratively updated while the distance between pixels and centroid is measured. The centroid value is updated via Eq.2.4.

Drawback for FCM algorithm is the sensitiveness to noise or outlier. The algorithm treats outlier or noise as a different cluster due to the large difference of value between the cluster center and the noise. FCM enforces the constraints $\sum_{i=1}^{c} \mu_{ik} = 1$ at which every pixel will be forced to belong to at least one cluster that has the nearest distance to the cluster centroid. FCM also suffers from singularity problem. Singularity in FCM occurs when one or more of the distance between pixel and cluster centroid is at the same point at any instance (Pal et al., 2005). The term singularity means a point where variable meets uncountable or infinite value. In FCM, the membership update function employs formula in Eq.2.3, where μ_{ik} is the membership of pixel i to cluster k and D_{ik} is the distance between pixel i to center k. When $D_{ij} = 0$, μ_{ik} become infinite, singularity occurs. As FCM enforces the principal of sum of total memberships to cluster centroid for a pixel is equal to one , FCM is very strong to force a pixel to assign to a cluster which will affect the overall result of the clustering.

FCM problem was solved by introducing PCM. In order to relax the restriction of FCM, the sum of the column or clusters' values must be equal to one, PCM was introduced by Krishnapuram and Keller (1993). PCM clustering algorithm is minimized by the objective function in Eq.2.5. PCM uses Eq.2.6 as membership update function and Eq.2.4 as centroid update function. During the process of FCM clustering, there is the tendency of one pixel having the same membership values between two clusters. PCM expresses membership values with formal constraint where the sum of column or cluster's to values of one $\sum_{i=1}^{c} \mu_{ik} = 1$. The idea of PCM is to enable the membership values to behave independently within the same cluster. Thus, the total sum of membership to clusters can be between zeros and one, $0 \leq \sum_{i=1}^{c} \mu_{ik} \leq 1$. PCM relaxes the constraint of FCM on the sum of column or cluster values must be equal to one. Another idea of PCM is that the distance of pixel to cluster center is inversely proportional to the influence on the membership. As pixel lies further apart from the cluster center, PCM tends to assign lower membership value to it. For noise, the pixel will have different

values from the cluster, therefore the membership of the noise pixel will be assigned to smaller membership values according to PCM algorithm. This, PCM helps noise reduction. Besides that, PCM also overcomes the needs for specifying the number of cluster. Through initiating a relatively big random figure, the number of clusters will be automatically identified when the objective function is minimized. PCM is robust to noise, one of the main concerns in medical imaging. Therefore, PCM is proposed to be used in this research due to its noise robustness feature.

However, PCM is very much sensitive to the initialization of cluster center v_i as compared to FCM. In PCM, the clusters do not have a lot of mobility, since each pixels will only related to one cluster at a time rather then all the clusters simultaneously. The generated output is greatly dependent on the initialization of cluster center. However, the freedom of pixel to act independently in the data set, produces weak cluster.

FGcM clustering algorithm proposed by Menard and Eboueya (2002); Menard et al. (2003) includes the Tsallis information (Tsallis, 2002). Tsallis information takes into account points that can be far from prototype and spaces out centroid and more than the classical objective function. FGcM allows one to deal better with inhomogeneity of signal intensity than FCM. In FGcM, the singularity problem as faced by FCM is solved. It incorporates Tsallis entropy which takes into account long range interactions. The possibility a pixel to be assigned to a cluster using entropy is higher. The objective function includes Bezdek objective function (James, 1981), Tsallis entropy and probabilistic constraint. The FGcM clustering algorithm converges to the original FCM when fuzzyness tends to 1 as the Tsallis entropy converges to Shannon entropy. Tsallis entropy has been used to replace Shannon entropy to embrace the full range of possibilities from a non-extensive setting (Menard et al., 2003), as presented by the more general Tsallis framework. The parameter m in Tsallis entropy is the measurement of the degree of the non-extensibility. Usually entropy is considered extensive, which means

its values are proportional to the system size. Tsallis statistics are based on generalized entropy. Its application deals with systems which have points with long range interactions when entropy becomes non-extensive. For short range interactions, it is interchangeable with the Gibs-Boltzman statistic and converges to Shannon entropy. The term in the FGcM objective function is the probabilistic constraint , where it can be derived using the EPI principle (Menard and Eboueya, 2002). Its justifies the imposed constraint on the FCM method and provides a method to derive the FCM objective function along with the probabilistic constraint.

However, FGcM suffers from the needs of knowing the number of cluster in advance before any clustering process is carried out. Number of cluster will effect the overall segmentation process. As FGcM evolved from FCM, the major problem of FCM such as singularity could be solved but it creates other new challenges. FGcM is computational expensive because of the complexness of the objective function.

Tsallis information also could be extended to PCM clustering algorithm which evolves to PGcM Menard and Eboueya (2002); Menard et al. (2003). PGcM considers long range interaction during the clustering algorithm. Certain membership values are assigned to the pixels eventhough they are a distance away from cluster center. FGcM deals better with intensity inhomogeneity, and it has the properties of PCM which is robust to noise or outlier. PGcM overcomes the need to know the number of clusters in advance. As to compare between PCM and PGcM, PGcM takes care of long range interaction where PCM does not. Drawback of PGcM is the sensitiveness of the initial cluster center and does not have a lot of mobility as in PCM.

Roullier et al. (2007) proposed to combine FGcM and PGcM together for image segmentation. By knowing the number of clusters presented in the image, FGcM is suitable to process the image. FGcM produces reasonable cluster centers at the end of the process. The cluster

centers are then used as the initial cluster center by PGcM for image segmentation. Actual image segmentation is done by PGcM because PGcM is robust to noise.

3.2 Incorporating spatial neighborhood relationship

Image intensity inhomogeneity is also known as bias field, intensity non-uniformity or shading. It can be observed as a smooth variation or changes of intensity across the image. There are several methods introduced to solve inhomogeneity of intensity. As discussed by Ahmed et al. (2002), modified FCM was used on medical image segmentation. They considered the objection function as the cause for pixel to be assigned to cluster center. They modified the objective function in such a way that the pixel can be assigned to the cluster if it is in the homogenous regions, in order to increase the robustness of FCM. Another technique to incorporate spatial information and also to reduce the computational time is by calculating the probability of neighboring pixels to the reference pixel (Chuang et al., 2006). Mandava et al. (2010) proposed a new spatial information function which includes the criteria of (Ahmed et al., 2002) and (Chuang et al., 2006). With the proposed method, the spatial constraint is adaptively taken into consideration by incorporating it in the membership function.

3.3 Kernel-based distance measurement

This section presents several improvements to the popular technique of FCM. Muller et al. (2001) proposed kernel function to replace Euclidean distance measurement in FCM. Kernel function was used to transform all pixels values into different dimensions in order to increase the separation boundaries between clusters. FCM employs the Euclidean distance measurement as the classification criteria. Euclidean distance measurement is very sensitive to outlier and noise which will lead to mis-clustering (Songcan and Daoqiang, 2004). Euclidean distance measure implicitly imposes the assumption of hyper-spherical clusters for data, therefore if