CLUSTER VALIDITY OF XIE AND BENI AND THE PARTITION COEFFICIENT INDEXES FOR FUZZY C-MEANS CLUSTERING

NOR AZRIN BT AHMAD MUSTAFFA

(2007286456)

THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR

BACHELOR OF COMPUTER SCIENCE (Hons.)

(CS230)

FACULTY OF COMPUTER AND MATHEMATICAL SCIENCES

UNIVERSITY OF TECHNOLOGY MARA

MALAYSIA

MAY 2010

ACKNOWLEDGEMENT

Syukur alhamdulillah, firstly I would like to express my thankful to ALLAH S.W.T for His mercy for giving me the strength and ability to finish up this project on time to fulfill the requirement for the Final Year Project (CSC 699) in order to obtain a Degree in Computer Science (Hons) at the Universiti Teknologi MARA.

I would to take this golden opportunity to dedicate my special thanks to my supervisor, Puan Rohana Bt Embong for encouraging, motivating and supporting me throughout the duration of completing this project. She has given me guidance and ideas to produce and complete this project within time. Without her cooperation and tolerant, I might face lots of problems. For Dr Nasiroh Omar and Pn Zaidah Ibrahim, thank you for all helps and ideas that have been given in order to complete my project.

I specially dedicate my love and thanks to my parents for their support and encouragement throughout my study at Universiti Teknologi MARA. They also provide spending money for me to complete this project. Last but not least, thank you to all my friends who directly or indirectly help me by giving ideas and support. Again, thank you so much to all of you.

MAY 2010 NOR AZRIN BT AHMAD MUSTAFFA

ABSTRACT

Under Image Processing, there is Image Segmentation. Image Segmentation is a subset of an expansive field of computer vision which deals with partition an image into meaningful regions with respect to a particular application. In particular, it is used to separate regions from the rest of the image, in order to recognize them as objects. In this project, we implement fuzzy c-means (FCM) clustering which is the technique of segmentation into mammographic images. Segmentation defines the boundary of the targeted object from its background in the images. This project focuses on suspected region that may contain breast anomalies such as masses and calcifications. These breast anomalies may be diagnosed as cancer by radiologists. Therefore, segmentation of mammographic images is an important phase in image analysis that can be further applied to other algorithms for specific tasks such as the detection and classification of breast anomalies. The implementation of FCM for the segmentation of mammographic images is by using Matlab. FCM is widely used technique in this regard but it requires the priori specification of the number of clusters. Therefore, this project is posed as one of optimization of a fuzzy cluster validity index. There are two validity measures in the context of fuzzy clustering that are being used which are Partition Coefficient and Xie and Beni index. We use C language to write down the cluster validity indexes.

TABLE OF CONTENTS