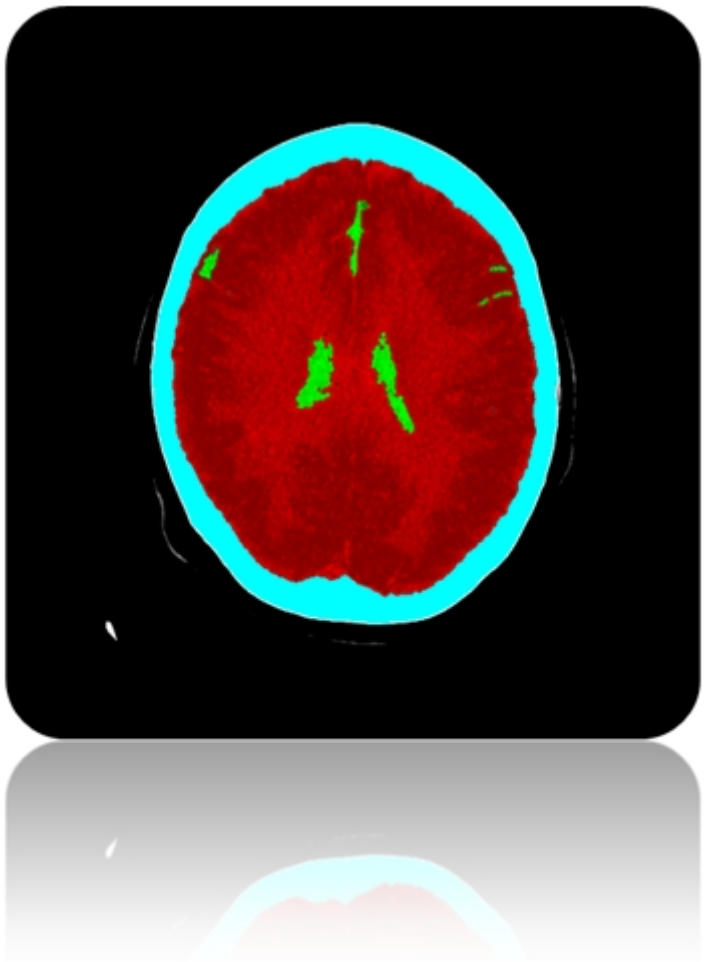
Aalborg University, Department of Electronic Systems
10th semester

Automatic segmentation on CT scans of human brain



Project group: 1024

Tamás Utasi

Supervisor:

Zheng-Hua Tan

Preface

This report has been written by project group 09gr1024 at the department of Electronics and Information Technology (ESN) at Aalborg University (AAU) during the 10th semester at (February 1st 2009 - August 30th, 2009).

The project is titled “Automatic segmentation on CT scans of human brain” and analyzes Fuzzy c-mean clustering technique applied on CT head images.

The report is divided into three main parts.

The first part (section 2) covers the general knowledge about medical imaging, segmentation and machine learning techniques.

The second part introduces the basics of the clustering algorithms.

The final part is about the implementation and the test results ended with a conclusion, the possible improvements, references and appendix.

Technology notice

This present document has been written using Microsoft Word. The graphics have been made with Corel PHOTO-PAINT X3.

Aalborg University 3th June, 2009

Signature

Abstract

The aim of the project is to make automatic segmentation on Computed axial Tomography (henceforward CT) scans of the human brain. The required labels are: Brain matter, Cerebral liquid, skull and background (included calcifications).

Magnetic Resonance Imaging (MRI) is better at differentiating soft tissue, but still there are some reasons why it worth to examine segmentation on CT images of the human brain.

- MRI can not used if the patient has any metallic equipment embedded in anywhere in the body,
- CT is more widely available in most hospitals,
- It is not suggested to exanimate a claustrophobic patient with MRI

In the current paper, Fuzzy c-mean - overlapping clustering – method is investigated with possible extension of the applied feature vectors.

Current results show FCM (even with the PDI extension) does not give optimal result, since it over classifies cerebrospinal liquid as brain matter.

Acknowledgments

This report covers a thesis project work. During the last five month I got numerous help from different sources. That is why I would like to thank all the people who have been involved in.

First of all I would like say thank you to my supervisor Zhen-Hua Tan, for his helpful contribution, consideration, and controlling my project..

I also would like to take the opportunity to say thank you to Lars Bo Larsen, coordinator of the Vision Graphics and Interactive System master program to provide me the opportunity to stay at AAU for one more semester.

Without important aid of the Hospital of my Home city I could not test my software, so I would like to testify my gratitude to Dr. Molnár Zsuzsa, to provide me the required CT scan series.

It was also really important to get some segmented images, that is why my thanks goes to Babos Magor, who performed the manually segmentation for me.

List of figures:

Chapter 1

Introduction

In this session, after a short preface, the main medical imaging technique will be presented. Finally the aim of this project will be defined.

1 Introduction

1.1 Context

Before advent of medical imaging, the examination of some disease like lung and breast cancer was hard. Nowadays X-ray, Ultra sound, MRI, CT and other medical imaging systems are commonly used and important part of daily practice.

It is important to note that CT, and in general medical imaging, is not the primary way to line up a diagnosis. It is a useful accessory technique to help the doctors with highlighting the modifications of a tissue – in general an area - but not the main contrivance.

CT imaging is involved widely in remedy. Here is a – not complete - list about the disease where CT inspection is used:

- Lung cancer (the most common kind of cancer)
- Breast cancer (the second common kind of cancer)
- Stroke
- Hypophysis micro/macro adenoma
- Cancer of eyes
- Hyperthyroidism of Adrenal
- Tumor of Larynux cartilage
- Wasting of nervous system caused by toxics
- Etc…

In general CT is good in differentiating bones from other tissues and also to localizing bleeding areas. Combined with different kind of radioactive emitter matters CT can be used for determining position and size of altered tissues. These tissues have different blood admission than the surrounding environment.

Examples are shown in Figure 1.1 and Figure 1.2.

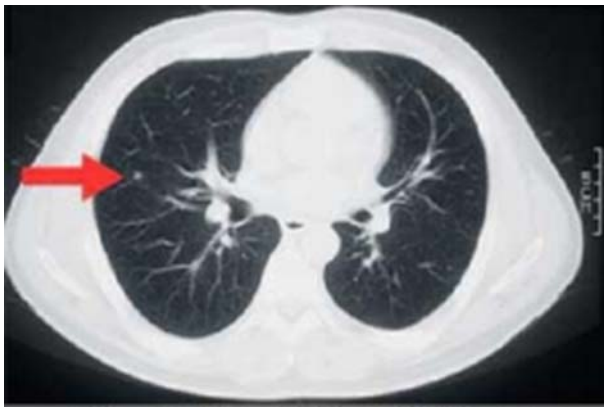


Figure 1.1 Small cancer bundle [10.]

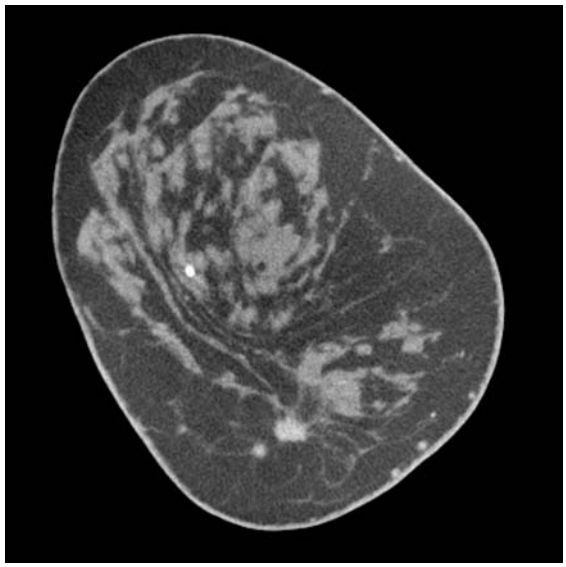


Figure 1.2 Altered region (white small dot left from the image center) [10.]

1.2 Objectives

The aims are:

- To apply machine learning algorithm on CT scans of human brain and give a label prediction about which sort of tissues are illustrated by the pixel.

- And as a byproduct to make a toolkit to improve the quality of the images, and a toolkit to manually segment the CT images.

To carry out this project it is important to investigate the currently used techniques in medical imaging to segment the different part of the human brain on CT scans. It is also recommended to get an outline on the field of machine learning about texture “understanding”.

Chapter 2

Pre analysis

In this chapter a brief introduction about the functional principle of CT will be provided. This part also aims at giving a quick list on the currently used techniques in segmentation techniques and place clustering in the field of machine learning.

2 Pre analysis

2.1 Medical imaging

2.1.1 Short history

Medical imaging and image processing is a big area in imaging and image processing. It is used in the daily remedy. From the electron microscopy to MRI, there is a wide variety of them. Medical imaging defined as techniques and processes used to create images of the human body for clinical purposes or medical science. In a wider sense it is part of biological imaging and it is used with radiology techniques.

In soft intellect, medical imaging is a way to create visual information about the target body or body parts without physically intruders into it. In hard intellect it is an inverse mathematic problem which means reproduce the inside structure of the given object from observations.

2.1.2 Imagine technologies

An extreme example of medical imaging technique is EEG (Electroencephalography – recording of electrical activity produced by the activated neurons within the brain in short period of time). EEG is not imaging technique in terms of daily imaging, but it is still useful, and is a widely used technique to visualize information about the parameters of brain. In further examples like ultrasound the reflected ultrasound waves, in the case of x-ray, the amount of x-ray pushed through show the inside structure of the organs. In the next part the most common used imaging techniques will be introduced.

2.1.2.1 Electron microscope

Electron microscope is a microscope which can magnify really small details. Electron microscope use electrons as the source of illumination, anchor with this technique an amassing opportunity to visually enlarge the target as good two billion times.

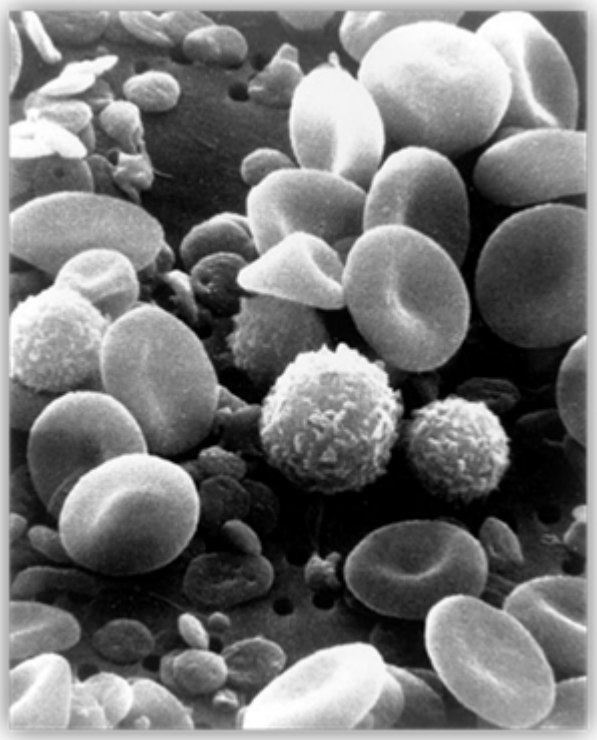


Figure 2.1 Electron microscope image of blood components [4.]

2.1.2.2 Radiographs (Projection radiography or Roentgen graphs)

The most common known and one of the oldest imaging technique used by the medical society is Radiographs. In fact radiograph is a 2D projection of the target object. The output is a cumulated image which means that the absorbed radiation is cumulating during it passes through the body. We cannot know the exact position of the tissue.

This technique suffers from some basic problems, since the beam source mostly point like.

Even with this problem it is really useful to detect fractures of bones or bundle of cancer.

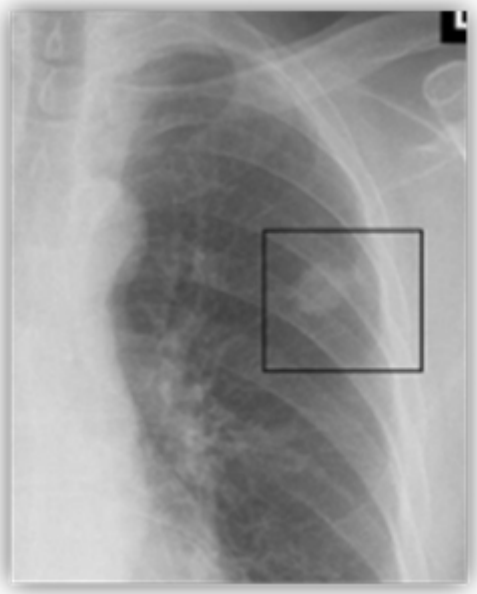


Figure 2.2 Cancer blob on x-ray image [10.]

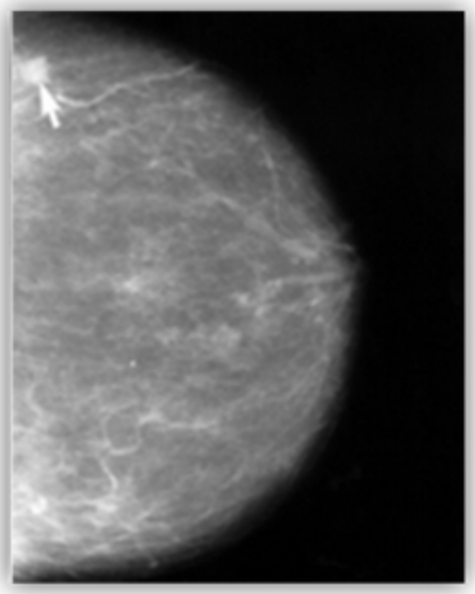


Figure 2.3 Breast cancer on x-ray image [9.]

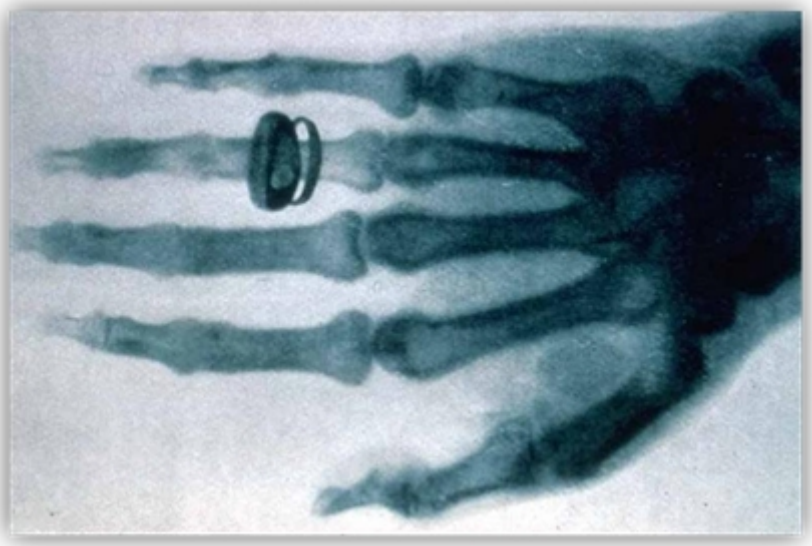


Figure 2.4. Hand of Alfred Kolikker © Wikipedia

Other terms like: X-ray, Film, Roentgen rays are also used.

2.1.2.3 Magnetic resonance imaging

A magnetic resonance imaging instrument (henceforward MRI, or as originally called "nuclear magnetic resonance (NMR) imaging") scanner uses powerful magnets to polarize and excite hydrogen nuclei in water molecules in human tissue, producing a detectable signal which is spatially encoded. Based on this signal a two dimensional image of a thin “slice” of the body is produced.

Unlike than X-ray, and that techniques in general which use ionizing radiation, there is no known side effect and therefore there is no limitation to the number of scans to which an individual can be subjected.

In MRI there is possibility to select which spin frequency will be excited, that is why it is possible to select which kind of tissue is going to be examined. It results excellent soft-tissue contrast achievable with MRI.

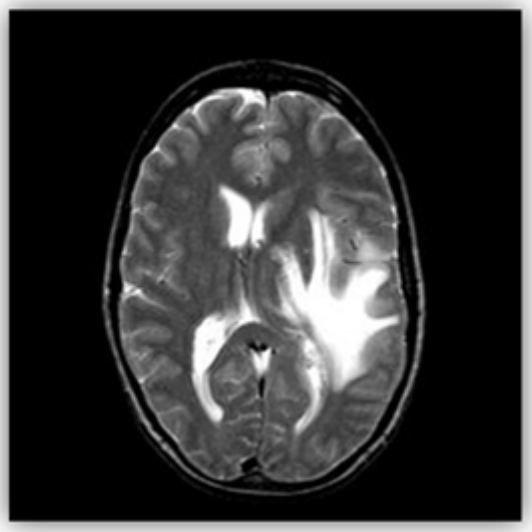


Figure 2.5 MRI brain image [5]

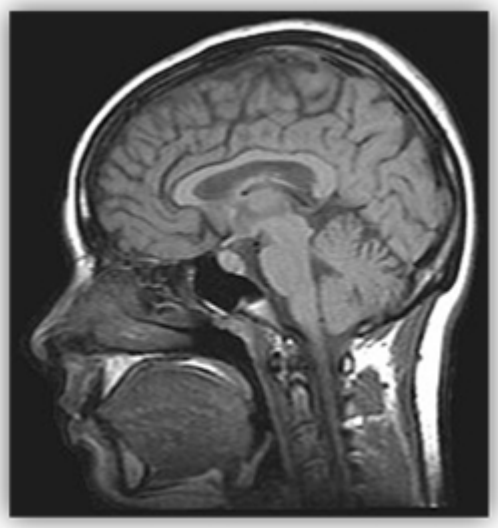


Figure 2.6 MRI brain image [6]

2.1.2.4 Nuclear medicine (SPECT/PET)

In nuclear medicine, pharmaceuticals labeled with radionuclide used to emit radiation which is detected. Later about the collected dataset an image is created. This image shows the density of radiation in the currently investigated volume, which denotes the activities in the area. (Activities require blood and other metabolic progressions that is why more radioactive matter is centralizing there.) The instrument which is used to detect the radiation is called gamma camera (it is also referred as Anger gamma camera).

Nuclear medical examination differ from most other imaging modalities in that the result mainly shows the physiological function of the member being investigated than again to traditional anatomical imaging such as CT or MRI.

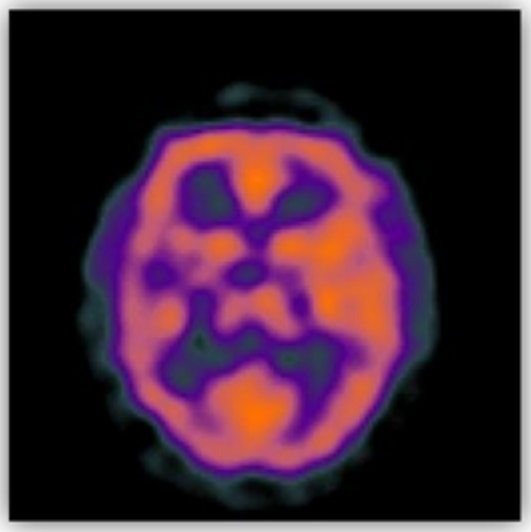


Figure 2.7 Typical SPECT image [5]

SPECT (Single Photon Emission Computed Tomography) and PET (Positron Emission Tomography) differs mainly in the involved physical phenomenon. While SPECT uses X-ray, PET uses gamma rays.

2.1.2.5 Computed axial Tomography

In the introduction chapter, the last “slice” making imaging technologies is CT. This is the mainly investigated imaging technology in this paper. For more details see the paragraph 2.2.

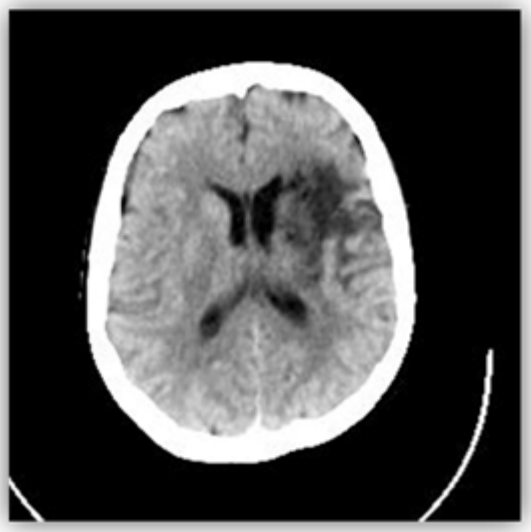


Figure 2.8 CT image of a brain [5]

2.1.2.6 Ultra sound (ultrasonography)

Ultra sound is cyclic sound pressure with a frequency greater then upper limit of human hearing (approximately 20 kHz). In medical imaging it used to

visualize muscles, tendons, and many internal organs. One of the biggest advantages of ultrasonography is its ability to produce real time image sequences. It is also relatively cheap and portable (especially compared with MRI or CT); moreover there is no known risk of using Ultra sound imaging. The frequencies can be anywhere between 2 and 18 MHz. The sound is focused either by the shape of the transducer (a lens in front of the transducer), or a complex set of control pulses from the ultrasound scanner machine. This focused ultra sound waves travels into the body and comes into focus at a desired depth. Some of the sound wave is reflected from the borders between different tissues. Specifically, sound is reflected anywhere there are density changes in the body: e.g. blood cells in blood plasma. Some of the reflections return to the transducer. This is collected and used to create different kind of images.

Four different modes of ultrasound are used in medical imaging. These are:

- **A-mode**: A single transducer scans a line through the body with the echoes plotted on screen as a function of depth.
- **B-mode**: A linear array of transducers simultaneously scans a plane through the body that can be viewed as a two-dimensional image on screen.
- **M-mode**: It gives the possibility to detect motion. In m-mode a rapid sequence of B-mode scans whose images follow each other in sequence on screen enables doctors to see and measure range of motion.
- **Doppler mode**: This mode makes use of the Doppler effect in measuring and visualizing blood flow

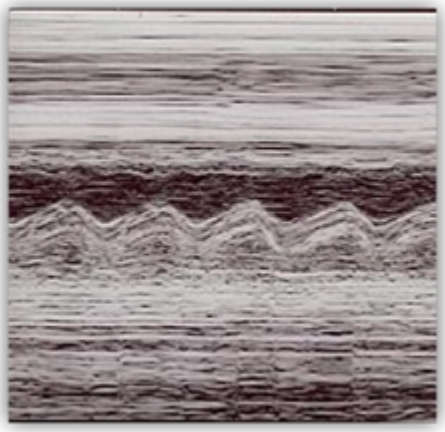


Figure 2.9. A-mode image. The periodical movement of the heart is clearly visible. [8]

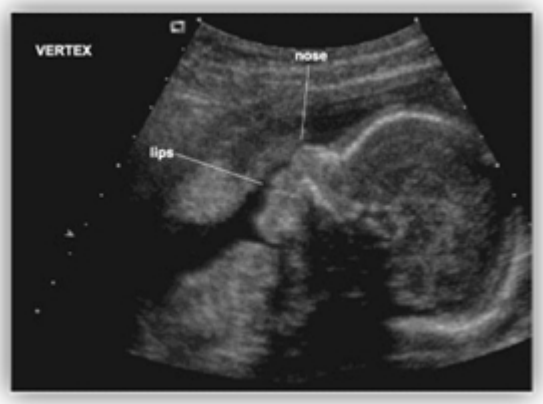


Figure 2.10. Image taken about pregnancy in B-mode. [8]

2.2 Functional principle of CT

2.2.1 Short introduction

The main principle of CT is really simple. The inside structure of an object is calculable if measurements of projections from different angles are given. The mathematical background was introduced by J. Radon in 1917. The first acting machine was built in the '70s by Allan McLeod Cormack (23th February 1924 – 7th of May 1998) and Sir Godfrey Newbold Hounsfield (28th August 1919 – 12th August 2004). In September 1971, CT scanning was introduced into medical practice with a successful scan on a cerebral cyst patient at Atkinson Morley Hospital in London. With this performance CT was the first imaging machine which could provide detailed structural information of internal three-dimensional anatomy of living creatures.

2.2.2 Hounsfield unit

The Hounsfield unit (HU) scale is a linear transformation of the linear attenuation coefficient of the original attenuation coefficient. The transformation is given by formulation:

$$\frac{\mu_X - \mu_{H_2O}}{\mu_{H_2O}} \times 1000 \qquad (1)$$

where μ_X the attenuation coefficient of matter X and μ_{H_2O} is the attenuation coefficient of distilled water. A change of one Hounsfield unit (HU) represents a change of 0.1% of the attenuation coefficient of water since the attenuation coefficient of air is nearly zero.

Substance	Hounsfield unit
Air	-1000
Fat	-120
Water	0
Muscle	+40
Contrast	+130
Bone	+400 and 400+

The above standards were chosen as they are universally available references and suited to the key application for which computed axial tomography was developed: imaging the inside anatomy of living creatures based on organized water structures.

The fix points of the scale are:

- Distilled water: 0
- Air: -1000
- The maximum possible value is 3000, but some of the manufacturers extend this range, to produce more detailed result.

Likely in projection radiography, x-ray beams are used but in case of CT - instead of a film – detectors are collecting the signals. Later, based on these measurements, a reconstructed image is created aided by computers. During creating a projection some problems are arising. These problems can be familiar from Rontgen graph.

2.2.3 Arising problems during creating a projection

Since the radioactive beam source is point likely, the mapping is suffer from projection.

2.2.3.1 Distortions

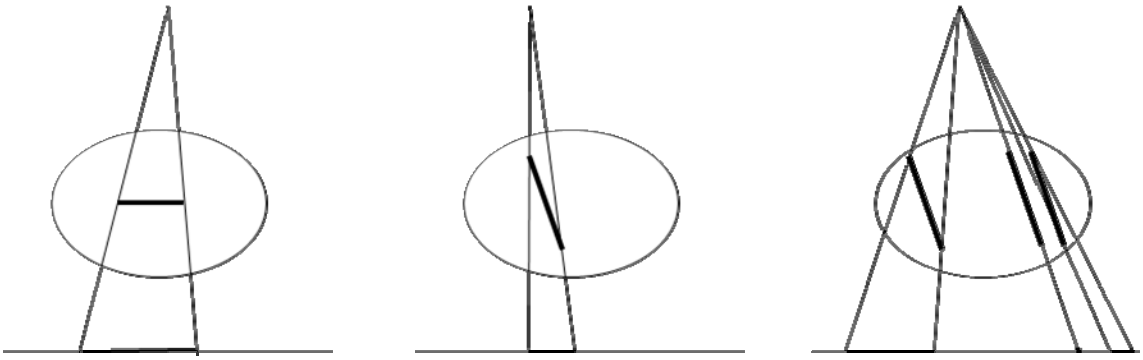
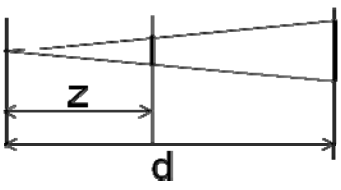


Figure 2.11. Three possible configurations illustrated from the endurable distortions.

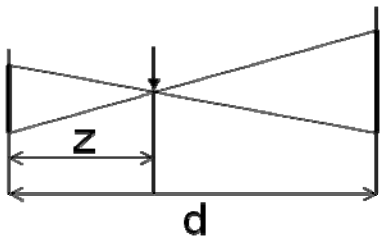
2.2.3.2 Enlarging

Basically there is two different kinds of enlarging effect. The enlarging of the object (Figure 2.12.) and the enlarging of the source (Figure 2.13.).



The enlarging is harmonic with: $\frac{d}{z}$

Figure 2.12. Enlarge of the object



The enlarging is harmonic with: $\frac{d-z}{z}$

Figure 2.13. Enlarge of the source

Where

- **z** denotes the distance between the source and the image plane and
- **d** denotes the distance between the source and the target object.

2.2.3.3 Density

There is a problem with the different tissue with different size and different density. If an internal organ has bigger size and less density it can still appear on the image like a smaller object with higher density value (Figure 2.14)

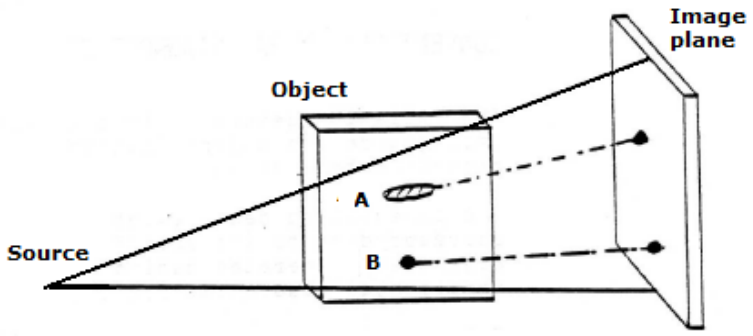


Figure 2.14. Example of mapping of tissues with different size and density

The object A has lower density value but bigger volume on the axes of the beam compare to object B which has higher density value and smaller coverage. Even with different conditions they have the same appearance on the image.

2.2.3.4 Mapping objects in different deepness

Object in different deepness are mapped with different size. This distortion is harmonic with the distance from the source (see 2.3.2).

2.2.4 Measured values

When a projection is created in a defined angle the measured signal is given by the next equitation:

$$I = I_0 e^{-\mu d}$$

Where

- I the measured signal (the leaving radiation from the object),
- I_0 the radiation which enters into the target,
- e is the base of natural logarithm, and finally
- μ radiation attenuation coefficient

μ is characteristic of the tissue and depends on the consistency of the matter and the spectrum of the Rontgen radiation.

On the x-ray image the higher leaving radiation appear like darker area. It follows that if the tissue has lower attenuation coefficient value, more radiation will leave the body and the defined picture point will be darker.

During the examination for each slice, more than one thousand projection created in different angles. Each projection radiation is detected by hundreds of sensors. This measured dataset called raw data.

2.2.5 Reconstruction

During the reconstruction the signals are transformed into a matrix. One item of the matrix is corresponding to the pixel of the result image. The image matrix of a typical CT machine has the dimension 512 x 512. Each pixel represents an amount of volume of the currently examined member. It is called voxel.

To reconstruct the internal structure of the investigated object more techniques are presented until today, starting from the exact mathematical reconstruction problem, until different kind of approximation methods. The exact solution method requires too much calculation power and time complexity. This and because approximation methods can produce close the same result, approximation methods are preferred. The most commonly used method is the *filtered back projection* [search some reference].

2.2.5.1 Calculating the HU values

After the back projection each item of the image matrix contains a value which denotes the linear attenuation coefficient of the voxel. In the next step the computer determines a new value according to the (1) equation.

2.2.6 Visualization of CT scans

With finishing the reconstruction phase the procedure of creating the CT image could be considered as finished as well, but it should not be displayed directly on a screen, since the produced image has at least 4000 different gray levels. The human vision system is not able to distinguish such numerous different intensities that is why some kinds of transformation are required.

To make it acceptable to human interpreters, a simple technique is used: it is called Volume Of Interesting windowing (VOI windowing). Instead of the whole Hounsfield spectrum, just a narrower range of it is displayed.

The transformation has two main parameters;

1. Center: This parameter defines the center of the window on the Hounsfield range.
2. Width: This parameter shows the biggest distance between the lowest and the highest interpreted HU value.

This window mapped into the range [0,255] which is ready to display. Under the lowest interpreted HU value the output intensities is defined as zero, as long as above of the highest end of the window the output intensities is defined as 255.

2.3 Segmentation techniques

Image segmentation is a process of partitioning the image into sets of pixels. Nowadays a wide range of segmentation toolkits are introduced in the field of image processing. The list can start from the simplest one, thresholding to more sophisticated algorithms like neural networks or other machine learning algorithms. The field of texture understanding is also growing.

Since segmentation of CT head images is not easy, few different approaches were introduced until today like statistical pattern recognition, morphological processing with thresholding, active contours, and clustering algorithms.

The challenges of segmenting CT head scan are increased with partial volume effects which impress the edges, produce low brain tissue contrast and fuse different objects within the same range of intensity.

However, wide varieties of segmentation techniques are available in image processing. The list below is created without claim of completeness.

2.3.1 Histogram based methods

The histogram of the image is calculated and then according to the peaks and valleys the pixels are grouped into clusters. This is a really efficient way of segmentation when the average of the object and the background pixel are separable.

2.3.2 Edge detection methods

Regions and its boundary are closely coherent. So giving the border or the area itself is analogous. The edge detection techniques are well-developed part of image processing.

2.3.3 Region growing methods

This technique is also used in the present project, like a partial development, to give a toolkit to technician to manually segment some CT scans. A basic version of region growing is when a seed point, a similarity degree and tolerance value are given. The regions are iteratively enlarged by comparing the unchecked neighbors. If the difference between the currently investigated pixel and its investigated neighbor according to the similarity degree is less than the tolerant value, then the neighbor considered as part of the region, and its neighbors are also investigated. One variant of this technique, proposed by Haralick and Shapiro [1] is based on pixel intensities. The mean and scatter of the region and the intensity of the investigated pixel is used to compute a test statistic. If the test statistic is small enough, then the pixel is added to the region, and the region's mean and scatter are updated. Otherwise, the pixel is rejected, and is used to form a new region.

2.3.4 Watershed transformation

The intuitive idea underlying this method comes from geography. Since any grayscale image can be considered as topographic surface: The intensity of the pixel is regarded as altitude of the point. Let us imagine the surface of this relief being immersed in water. Holes are created in local minima's of the surface. Water fills up the dark areas "the basins" starting at these local minima. Where waters coming from different basins meet, dams are built. When the water level has reached the highest peak in the landscape, the process is stopped. As a result, the landscape is partitioned into regions or basins separated by dams, called watershed lines or simply watersheds.

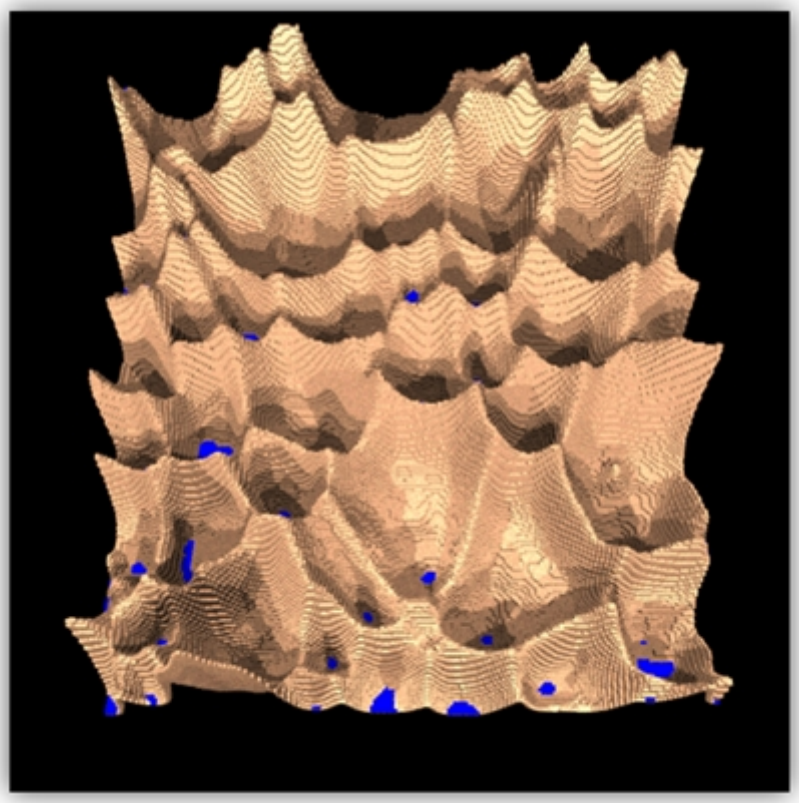


Figure 2.15. The pursuit of watershed algorithm illustrated. [2]

2.3.5 Neural network segmentation

Neural Network segmentation relies on processing small areas of an image using an artificial neural network or a set of neural networks. After such processing, the decision-making mechanism marks the pixels of the image accordingly to the category recognized by the neural network. A type of network designed especially for image processing is Kohonen map, also referenced as SOM Self Organizing Map). It is a rich source in the literature of image processing.

2.3.6 Semi-automatic segmentations

In this approach the user outlines the region on some way, typically using the mouse, and after an algorithm - initialized with the user input - applied on the image which fits the path on the image used some kind of similarity measurement.

2.3.7 Active contours (or snakes)

Active contour also called snake tries to minimize an energy associated to the current contour as a sum of an internal and external energy.

2.3.8 Clustering segmentation

Clustering is an iterative technique used to partition an image into segments. The basic algorithm:

1. Set up k cluster center (typically it is a vector). The initialization can happen randomly or based on some heuristic.
2. Mark each pixel on the image with that cluster which minimizes the variance between the pixel and the cluster center.
3. Update the center of the clusters.
4. Count objective function. If the function is converged (e.g. no change in the marking of the pixels) then stops, otherwise repeats 2. and 3.

The variance can be the squared or absolute difference between a pixel and a cluster center. The difference is typically based on pixel color, intensity, texture, and location, or a weighted combination of these factors. *K* can be selected manually, randomly, or by a heuristic.

2.4 Machine learning

Machine learning is a scientific discipline which deals with design and development of algorithms that allows computers to learn based on given data set. A major focus of machine learning research is to automatically learn to recognize complex patterns and make intelligent decisions based on data. In general computer programming, the exact way to solve the problem is implemented. In machine learning the implemented code try to discover the way, how to solve the given problem.

Machine learning algorithms are collected into the following main groups:

2.4.1 Supervised learning

A set of (input) measurements and their outputs are given. This is called training data set. In case the output is a continuous function, this process is called regression and then when the outputs are labels (or class names) it is called classification. During the learning phase the labeled data provide information about the decision error. Based on this error the system tries to give an improvement of it (reducing the error).

In general, supervised learning generates a global model of how to solve the given problem (the mapping between the input and output). However, in some cases the solution implemented as set of local models of the problem. The input usually is given like a vector, which contains a way of description of the object. This vector is called feature vector. The quality of the system strongly depends on the representation of the object. It is also important that the input has to be normalized on some way. Depending on the desired output function, different learning methods should be used. For example: to learn a continuous function a decision tree is unusable, again a neural network is more efficient.

2.4.2 Unsupervised learning

Unsupervised learning is distinguished from supervised and re-informed learning methods in that the learner is given only unlabeled data. The learner method has to find the inside structure of the problem in itself. One form of unsupervised learning is clustering (see more details in the analysis part).

2.4.3 Reinforced learning

Reinforced learning is almost like supervised learning. In this case during the learning phase, instead of the desired output only positive or negative feedbacks are given.

Finally a practical example of machine learning: Let's look the process of learning a language. Supervised learning: given some exercises. During the learning progress the examples are evaluated by the learner and corrected by the "teacher". The mistakes are listed, so the learner has the opportunity to learn from its mistakes. In case of unsupervised learning a rule book given and it is the trust of the learner to pick up the knowledge from the book. In case of reinforced learning the exercises are given again, and the "teacher" is correcting the example too, but instead of detailed feedback, only a positive or negative answer is given to the learner.

Chapter 3

Analysis

This chapter analyses the theorems used.

3 Analysis

3.1 Clustering in general

Clustering is one of the most commonly used unsupervised learning algorithms. Like in unsupervised learning engineering the basic conception is to explore the hidden structure of the problem on unlabeled data. During the learning phase, the examples are grouped into one set (cluster), which are more similar with each other than with the other examples.

So, the goal of clustering is to determine the intrinsic alignment in a set of unlabeled data.

The main requirements with clustering in general are:

- Scalability
- dealing with different types of attributes
- discovering clusters with arbitrary shape
- minimal requirements for domain knowledge to determine input parameters
- ability to deal with noise and outliers
- insensitivity to order of input records
- high dimensionality

There are a number of problems with clustering. Some of them are:

- current clustering techniques do not address all the requirements adequately (and concurrently)
- dealing with large number of dimensions and large number of data items can be problematic because of time complexity
- the effectiveness of the method depends on the definition of "distance" (for distance-based clustering)
- if an *obvious* distance measure doesn't exist we must "define" it, which is not always easy, especially in multi-dimensional spaces
- the result of the clustering algorithm (that in many cases can be arbitrary itself) can be interpreted in different ways

Clustering algorithms may be classified into the next four clusters:

1. exclusive clustering
2. overlapping clustering
3. hierarchical clustering and
4. probabilistic clustering

Exclusive clustering does not enable to a data to belong to several cluster, only one cluster can include the data, until overlapping clustering do the opposite. It allows the data to belong to multiple clusters in the same time. In this case the data will get the label of the most determining cluster (which has the highest possibility). Hierarchical clustering is based on merging the two closest clusters. Initially, every data is labeled as cluster. The probabilistic clustering assumes that the data are produced by a mixture of N multivariate of Gaussians.

K-means is an example of exclusive clustering algorithm, Fuzzy c-mean is a typical example of overlapping clustering. Hierarchical clustering is obvious and finally Mixture of Gaussian is a probabilistic algorithm.

Distance measurement is an important part of the realization. If the components of the data instance vectors are all in the same physical units then it is possible that the simple Euclidean distance metric is sufficient to successfully group similar data instances. However, even in this case the Euclidean distance can sometimes be misleading. Different scaling can lead to different results.

3.2 Fuzzy c-mean (FCM)

Fuzzy c-mean, as it said in the previous part, is an overlapping clustering method, which means that each piece of data could belong to two or even to all of the clusters. This method was developed by Dunn in 1973 and improved by Bezdek in 1981. It is frequently used in pattern recognition. Fuzzy c-mean based on minimization of the following objective function:

$$Minimize: FCM = \sum_{k=1}^{c}\sum_{i=1}^{N}\left(u_{ik}\right)^{m} d_{ik}^{2}\left(x_i, p_k\right) \qquad (2)$$

$$Subject\ to: \sum_{k=1}^{c} u_{ik} = 1, \forall i \in \{0..N-1\} \qquad (3)$$

c is the number of clusters, N number of the data (in current case the number of the pixels on the image), and m > 1 a real number which controlling the fuzzy property of the algorithm. u_{ik} the degree of membership of i[th] data in k[th] cluster. x_i the feature vector of the i[th] pixel and finally p_k is the center of the k[th] cluster.

Fuzzy partitioning is going through an iterative optimization of the objective function, with updating the membership values and the cluster centers as follows:

$$u_{ik}^{t+1} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}^2}{d_{jk}^2} \right)^{\frac{1}{(m-1)}}} \tag{4}$$

$$p_k^{t+1} = \frac{\sum_{i=0}^{N-1} u_{ik}^m x_i}{\sum_{i=0}^{N-1} u_{ik}^m} \tag{5}$$

In current implementation the distance metric is defined as

$$d_{ik}^2 \left(x_i, p_k \right) = \left\| x_i - p_k \right\|_A^2 = \left(x_i - p_k \right)^T A \left(x_i - p_k \right) \tag{6}$$

where A is positive, definite matrix. A initialized as the identity matrix, but it is updated in each iteration according to the next formula:

$$A_k^{t+1} = \frac{\sum_{i=0}^{N-1} u_{ik} \left(x_i - p_k^{t+1} \right) \left(x_i - p_k^{t+1} \right)}{\sum_{i=0}^{N-1} u_{ik}} \tag{7}$$

A is functioning as covariance matrix. So in practice, it means that the distance measurement used is the Mahalanobis distance.

Mahalanobis distance is based on correlations between variables. Different patterns can be identified and analyzed. It is a useful way of determining similarity of an unknown sample set to a known one. It differs from Euclidean distance in that it takes into account the correlations of the data set and is scale-invariant.

However, FCM has a drawback. It prefers the clusters with large size and/or diameter. (Size defined as the size of the population, and diameter as the diameter of the hyper-sphere which contains the entire cluster.)

3.3 Population-Diameter Independent algorithm

To compensate for these shortcomings in FCM, Shihab [3] proposed the Population-Diameter Independent algorithm (PDI). It has a new objective function in which each cluster contribution is normalized. ρ_k is the new tag in the formulations and be up to normalize the contribution of the k[th] cluster. Each clusters normalization value initialized with $\frac{1}{c}$. During the iteration period each normalization value is updated according to (8). The update formulation of the membership value is also changed to reflect the impact of ρ_k (9).

So the final form of the equitation is:

$$Minimize: FCM = \sum_{k=1}^{c} \frac{1}{\rho_k^r} \sum_{i=1}^{N} \left(u_{ik}\right)^m d_{ik}^2 \left(x_i, p_k\right) \tag{8}$$

$$Subject\ to: \sum_{k=1}^{c} u_{ik} = 1, \forall i \in \{0..N-1\} \tag{9}$$

$$u_{ik}^{t+1} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}^2}{d_{jk}^2} \right)^{\frac{1}{(m-1)}}} \tag{10}$$

$$\rho_k^{t+1} = \frac{\sum_{i=0}^{N-1} u_{ik}^m x_i}{\sum_{i=0}^{N-1} u_{ik}^m} \tag{11}$$

$$p_k^{t+1} = \frac{\sum_{i=0}^{N-1} u_{ik}^m x_i}{\sum_{i=0}^{N-1} u_{ik}^m} \tag{12}$$

$$A_k^{t+1} = \frac{\sum_{i=0}^{N-1} u_{ik} \left(x_i - p_k^{t+1}\right)\left(x_i - p_k^{t+1}\right)}{\sum_{i=0}^{N-1} u_{ik}} \tag{13}$$

The steps of the algorithm are when c and m are given:

1. Initialize A with the identity matrix. The mean vector (p_k) pre with fixed numbers (at least the first tree mean vector). The values come from empirical observation. With this step the result are reproducible, which is important.
2. Update the membership values and the covariance matrix according to 4-5, 7.
3. Compare the change in the membership value between times t and t+1. If the distance less than a pre-defined epsilon then stop. Otherwise repeat step 2.

3.4 The used feature vector

The first implementation of this project is based on [4]. Alexandra Lauric and Sarah Frisken in their work used a simple feature vector based on the intensity of the pixel at position *(x,y)* and its eight neighbors average.

Their result showed the PDI version of FCM made mistakes, and mostly, the brain matter around the cerebral liquid were wrongly labeled as cerebral liquid.

In this work some extra features added to increase the outcome of the FCM algorithm.

3.4.1 Histogram Moments

A n^{th} moments of a probability variable (here pixel value) is defined as:

$$\sum_{i=1}^{N} \left(x_i - m\right)^n \, p\left(x_i\right)$$

In general the central moments are good in characterizing the textures.

E.g.: The second central moment is good in representing the contrast of the image, the third central moment is featuring the shape of the histogram (symmetry) and finally the fourth central moment shows the flatness of the histogram. The drawback of central moments is that they do not describe any geometric relation between pixels.

In the present paper only the first moment is used.

3.4.2 Features based on co-occurrence matrix

The Grey Level Co-occurrence Matrix (GLCM) is a pixel-based well known statistic used for texture analysis, because it provides some information like the texture contrast, homogeneity, entropy, energy, correlation. It computes, for each possible pair of grey levels (l, m), the number of pairs of pixels, having intensities l and m which are situated from each other at a distance given by a specified displacement vector (dx, dy).

The gray level co-occurrence matrix is describing the correlation between pixels in the same direction and distance and it is defined as:

$$C(l,m|\Delta i,\Delta j) = \begin{matrix} occurence\{I(i,j)=l \,\&+\, I(i+\Delta i, j+\Delta j)=m\}+ \\ occurence\{I(i,j)=l \,\&+\, I(i-\Delta i, j-\Delta j)=m\} \end{matrix}$$

One item of the $C(l,m|\Delta i,\Delta j)$ defines the probability that if the pixel on position (*x,y*) has intensity l than the pixel in the direction defined by $(\Delta i,\Delta j)$ has intensity m.

The next features are derived from GLCO (Gray Level Co-occurrence Matrix) and are used in the present project:

Energy: $E(\Delta i,\Delta j) = \sum_{l=0}^{N}\sum_{m=0}^{N} C(l,m|\Delta i,\Delta j)^2$

Entropy: $H(\Delta i,\Delta j) = \sum_{l=0}^{N}\sum_{m=0}^{N} C(l,m|\Delta i,\Delta j)\log C(l,m|\Delta i,\Delta j)$

Chapter 4

Design

.

4 Design

4.1 Quality enhancement

Like most of the case in image processing, the assumption of good results is the correctly prepared point of origin. For the better result in the segmentation quality enhancement is performed on the CT scans.

Digital Imaging and Communications in Medicine (henceforward DICOM) is a collection of standard for handling, storing and transferring medical data. It defines the file formats and network transfer protocol as well. The aim of engineers of DICOM was to create a system which can handle and integrate data from different kinds of scanner (from different brands), servers, workstations and network hardware's. The DICOM contains some pre-defined standard and also some free headers. One DICOM file can store one picture, a whole scan series or even animation. It is possible to compress the data with different standards like JPEG, LZW or RLE. DICOM groups information into data sets. That means the file of a CT-image of the head, for example, contains the patient's name, age and other personal information, anchor on this way the image can never be separated from this information by mistake.

DICOM store pixel information in 32 bits. In a standard grayscale image 8 bit is enough to encode the gray level intensity (256 different gradiations). Figure 4.1 shows the result of the simplest mapping from the DICOM gray level range and the range [0,255]. As it easy to see this is not the best way since the area of the brain appears one, continuous gray region. It is true there is still some texture which is perceptible, but it is inefficient to separate the different parts of the brain from each other. Figure 4.3 shows the linear stretching of image a. It is not successful since the whole area appears to be white. On Figure 4.5 the histogram and the estimated histogram of input image plotted. It is apparent there are two relative high peaks (in order from left: the background and the bones). It is also easy to see, the information about the brain matter (on the histogram it is supposed to be between the two highest peaks) just vanished.

For good starting point it is important to perform a wiser conversion from DICOM to standard gray scale image. The same technique is used to convert from 32 bit grayscale image to 8 bit grayscale image like in case of visualization (paragraph 2.2.6).

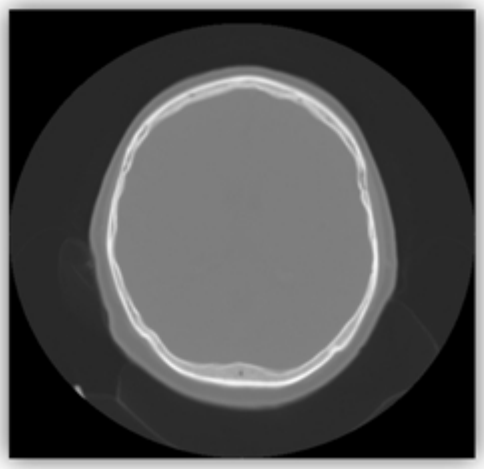


Figure 4.1

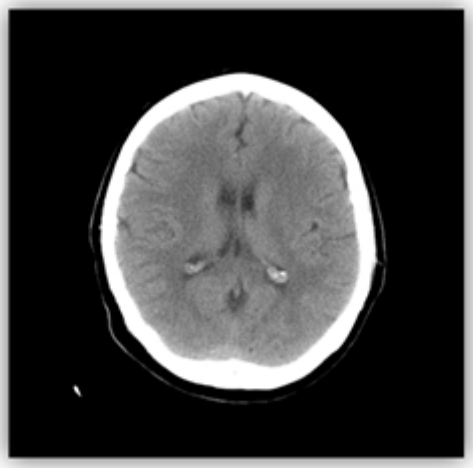


Figure 4.2

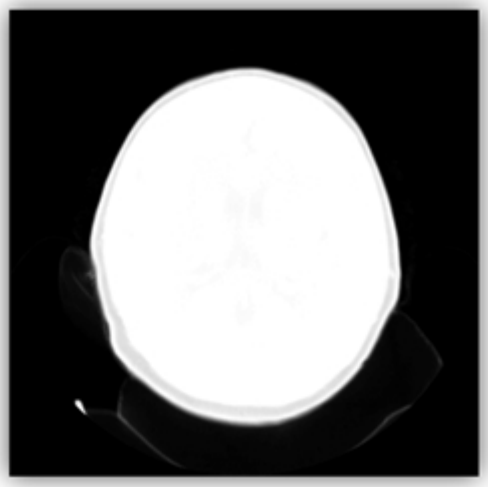


Figure 4.3

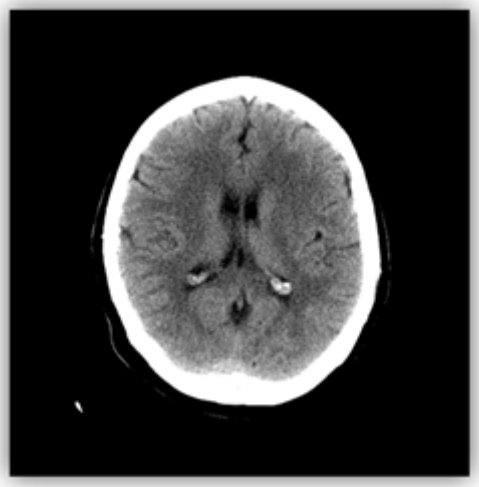


Figure 4.4

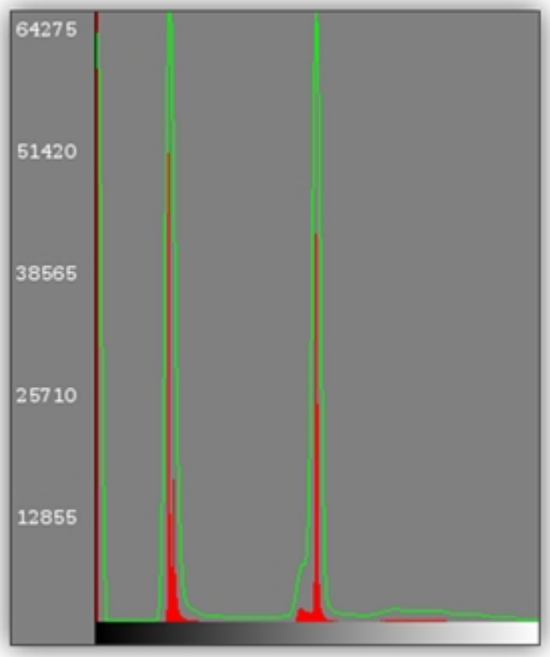


Figure 4.5

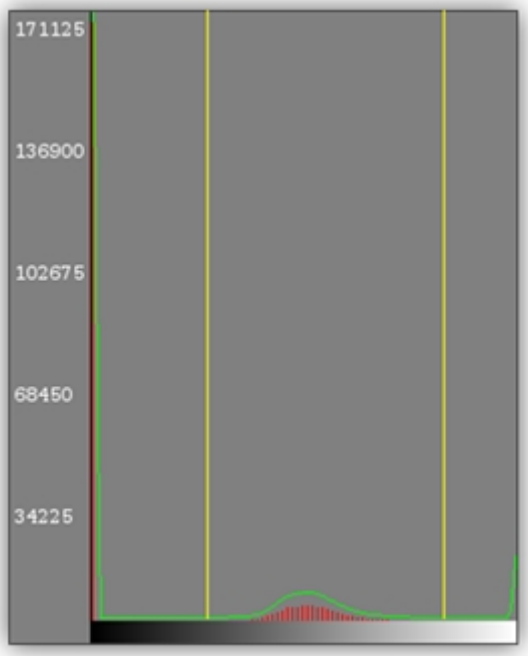


Figure 4.6

If the initial image is produced with VOI windowing, then the quality is still open to improvement.

To revise the nature of the CT scan the estimation of the histogram is counted. This step is advised since histogram of the image is not continuous Figure 4.9. For this reason a simple KDE (Kernel Density Estimation) method implemented.

In each pixel the Gaussian function

$$f(x) = a * e^{-\frac{(x-\hat{x})^2}{2c^2}}$$

placed weighted with the probability of the current pixel.

The Gaussian function parameters are the next:

1. $\hat{x}$ (correspond with the center of the function): the current intensity,
2. a (the height): is the probability of the intensity, the weighting is done with this parameter,
3. c (width): set up to 0.5, which effect the width of the “bell curve” is approximately 5 unit (pixel) in both direction.

Like a result the estimated (smoothed) histogram of the image is given.

Figure 4.7. Input image

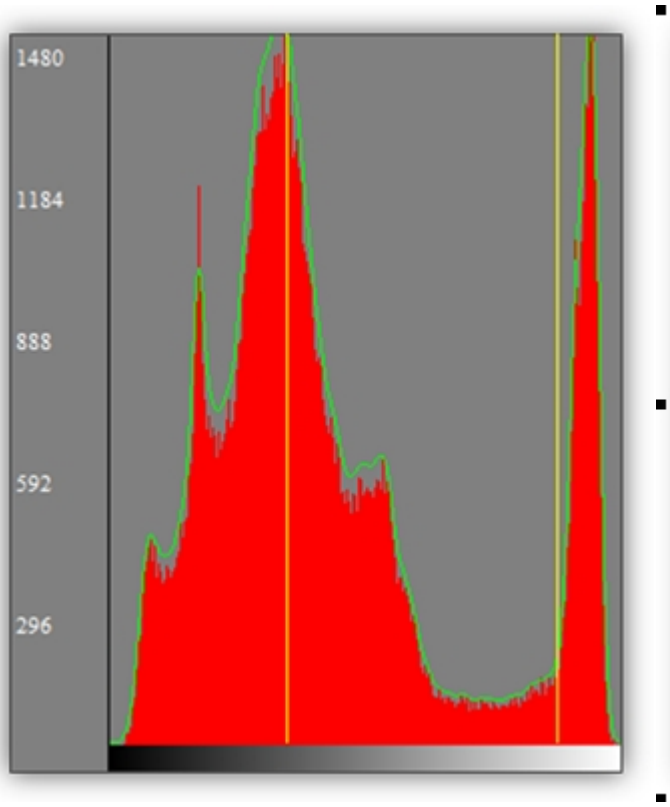


Figure 4.8. Real (red) and estimated histogram (green)

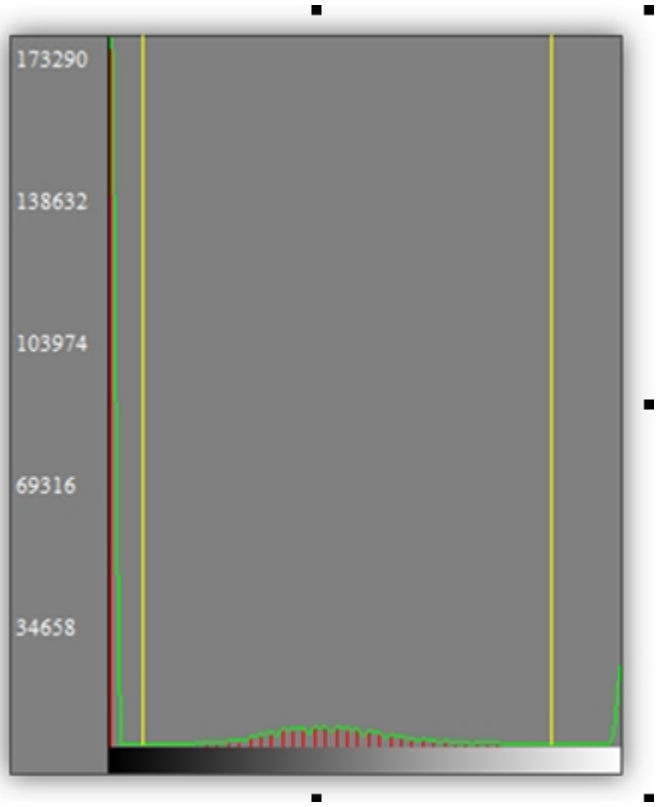


Figure 4.9. a histogram of a pre-enhanced CT Image

Let's consider the estimated histogram as a continuous function (in terms of mathematics it is still a discrete function) and let's search for the two highest peaks. Let's notify with Ba_{Max} the first peak from left (this is corresponding to the background) and Bo_{Max} the first peak from right (this is corresponding to the bones). After the positions Ba_{Max} and Bo_{Max} determined, the algorithm is starting to climb down inside of the two highest peaks. If the linear stretching would perform with considered Ba_{Max} like the minimum and Bo_{Max} like the higher threshold value, then the stretching would not be so efficient. When the climbing down one time reaches a dale (it can be even small), it stops climbing down and determines the new threshold value for the linear stretching.

4.2 *Quality measurement of the results*

The correctness of the segmentation is indispensable. It is required to measure on one way the accuracy of the algorithm and compare with the different results of method. For this reason some reference images – original and its corresponding segmented version – are required, since there is no easily accessible database a tool required to help the doctor to segment some enhanced CT scans.

4.2.1 Providing tool for manual segmentation

The basic idea is really simple. An image is displayed. The user is asked to point – by clicking – on the image to different kind of tissues. Every time a label is previously activated. When the user clicks on the image the pixel and its "adequate" connected neighbors will get the ID of the currently segmented tissue. The "adequate" connected neighbors selected according to the following steps:

- the gray level of the pixel is memorized
- its direct eight neighbors are investigated. If difference between the average of the eight neighbors of the currently investigated pixel and the selected pixel is less than a threshold value, then the pixel and it's position get the same label ID
- its own eight neighbors is going to be investigated as well
- the algorithm is going to stop when there is no more connected pixel to investigate.

4.2.2 Compare the results

For comparison, two commonly used index method are applied, by names:

- the confusion matrix and
- the precision with recall

In the confusion matrix, in each line, each cell indicates the percentage of classified a label (the label in the title of the line) as another label (the label in the title of the row). In the main diagonal cells, the percentages of the identical labels are indicated, which means the percentage of correctly determined cluster label is in these positions. The outside of this cell shows the delusions, when a pixel with a given label gets a different, false label. In case the sample and the currently compared image are the same, the confusion matrix has to be the identity matrix.

The other index method could be counted from the confusion matrix. Precision and recall are two commonly used statistical index-number. Precision and recall are defined in the field of information retrieval and statistical classification.

In terms of information retrieval, precision is defined as the number of the relevant documents retrieved by a search divided by the total number of documents retrieved by the search and recall is defined as the total number of relevant documents retrieved by search divided the total number of relevant document. So a perfect score (100%) in precision means that the result returned by the search was relevant (but does not say anything about whether the search returned with all of the relevant documents). In the other hand a perfect recall score (100%) means that all relevant documents were retrieved by the search (but says nothing about how many irrelevant documents were also retrieved).

In a statistical classification scenario, precision is defined as a number of items correctly classified belonging to the positive class (positive true) divided by the total number of elements labeled belonging to the positive class (included the false positive items also). Recall is defined as the number of true positives divided by the total number of elements that indeed belongs to the positive class.

In the current paper, precision and recall is defined as below:

		Correct labeling			
		Brain matter	Cerebrospinal fluid	Skull	Calcifications
Obtained result	Brain matter	$Tp1$	$Fn2_1$		
	Cerebrospinal fluid	$Fp2_1$	$Tp2$	$Fp2_3$	$Fp2_4$
	Skull		$Fn2_3$	$Tp3$	
	Calcifications		$Fn2_4$		$Tp4$

Figure 4.10. Precision and recall of label 2

Precision of class nr.2 (Cerebrospinal fluid) defined by the equation

$$Precision_2 = \frac{Tp2}{Tp2 + \sum Fp2_i} \qquad i = 1,\ldots,4 \quad i \neq 2 \qquad (14)$$

Along the same line with Recall defined as:

$$Recall_2 = \frac{Tp2}{Tp2 + \sum Fn2_i} \qquad i = 1,\ldots,4 \quad i \neq 2 \qquad (15)$$

The precision of label one, three and four are defined analogously.

Chapter 5

Implementation and Test

In this part, the condition of the implementation will be described.

5 Platform and development tools

The software is written in C++, developed on these different platforms: Ubuntu, Windows® Vista and Windows® XP. To compile and build the project Code::Blocks, an open source cross platform IDE with MinGW compiler were used. Theoretically the project is ready to build on Mac as well, but it was never tested before.

In the early state of the project OpenCV (Open Computer Vision library) were used. But because OpenCV has serious limitation in GUI (Graphical User Interface) building, and it was mostly used to access the image data, it has been replaced by wxWidget. wxWidget is a widely used cross platform library to create GUIs on wide palette of platforms. According to the official website of wxWidget, it supports the following platforms: Win32 (9x, NT, Xp, 2003, Vista), Win64 (Xp, 2003, Vista), Linux (kernel 2.4 and 2.6), NetBSD, FreeBSD, OpenBSD, Solaris, HP-UX, AIX, OS.

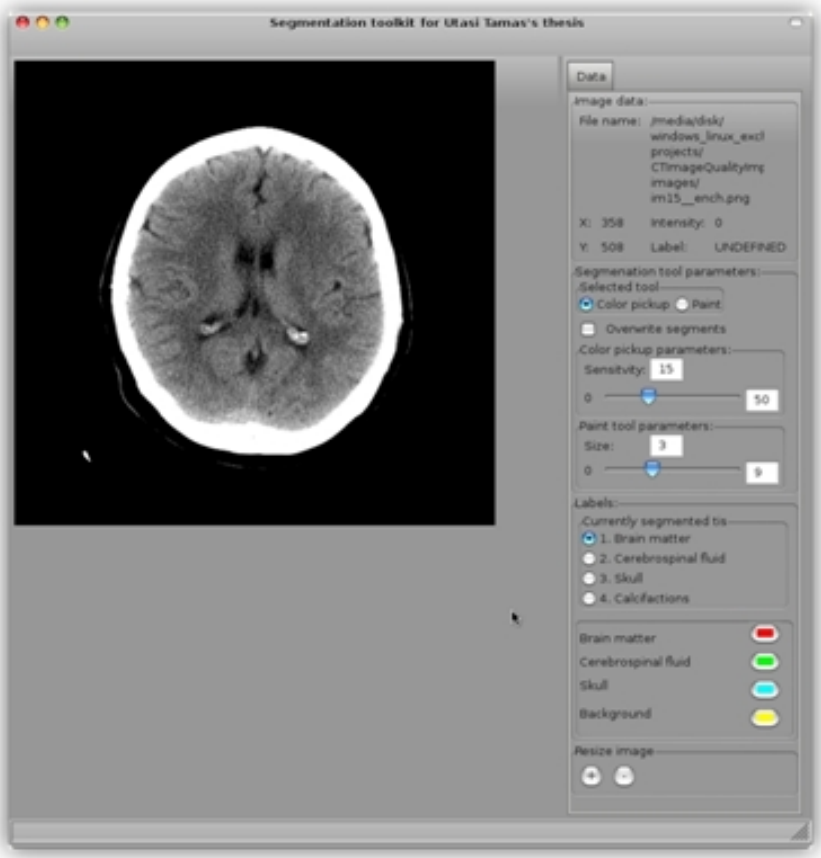


Figure 5.1. The GUI of the segmentation toolkit

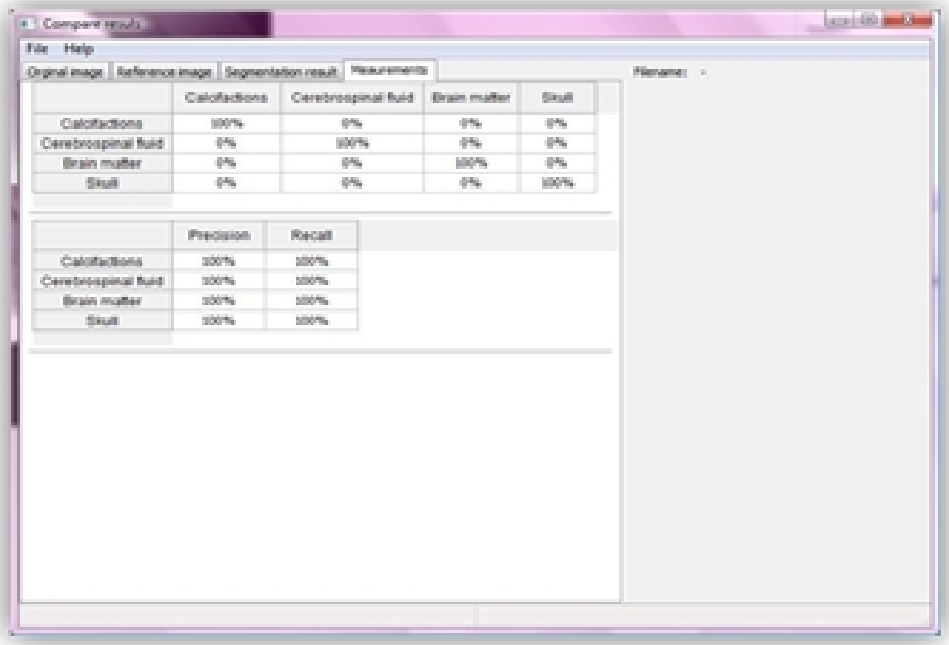


Figure 5.2. Screenshot of a theoreticaly perfect match.

	Calcifications	Cerebrospinal fluid	Brain matter	Skull
Calcifications	99.5%	0%	0%	0,542%
Cerebrospinal fluid	0%	1.61%	98.4%	0%
Brain matter	0.0146%	0%	97.6%	0.0961%
Skull	0.00976%	0%	9.91%	96.1%

	Precision	Recall
Calcifications	99.5%	100%
Cerebrospinal fluid	1.61%	100%
Brain matter	99.9%	49.4.4%
Skull	96.1%	99.3%

The recorded scores on test image Nr. 1. with the original version.

	Calcifications	Cerebrospinal fluid	Brain matter	Skull
Calcifications	99.6%	0%	0%	0,405%
Cerebrospinal fluid	38.8%	2.61%	58.6%	0%
Brain matter	2.39%	0%	97.6%	0.012%
Skull	5.99%	0%	0.0787%	9.39%

	Precision	Recall
Calcifications	99.6%	67.9%
Cerebrospinal fluid	2.61%	100%
Brain matter	97.6%	62.4%
Skull	93.9%	99.6%

The recorded scores on test image Nr. 2. with the original version.

	Calcifications	Cerebrospinal fluid	Brain matter	Skull
Calcifications	99.8%	0%	0%	0,221%
Cerebrospinal fluid	0.317%	6.17%	93.5%	0%
Brain matter	0%	0%	100 %	0.00656%
Skull	0%	0%	9.74%	90.3%

	Precision	Recall
Calcifications	99.8%	99.7%
Cerebrospinal fluid	6.17%	100%
Brain matter	100 %	49.2%
Skull	90.3%	99.7%

The recorded scores on test image Nr. 3. with the original version.

	Calcifications	Cerebrospinal fluid	Brain matter	Skull
Calcifications	99.9%	0%	0%	0,0695%
Cerebrospinal fluid	0%	15.3%	82.8%	0%
Brain matter	0.0764%	0%	99.9 %	0%
Skull	0%	0%	11.8%	88.2%

	Precision	Recall
Calcifications	99.9%	98.1%
Cerebrospinal fluid	15.3%	100%
Brain matter	99.9%	51.4%
Skull	88.2%	99.9%

The recorded scores on test image 4. with the original version.

	Calcifications	Cerebrospinal fluid	Brain matter	Skull
Calcifications	99.5%	0%	0%	0,542%
Cerebrospinal fluid	0 %	20.1%	70.2%	6.7%
Brain matter	0.0561%	2.325%	99.1%	0%
Skull	0%	0%	11.2%	88.8%

	Precision	Recall
Calcifications	99.5%	98%
Cerebrospinal fluid	20.1%	96%
Brain matter	99.1%	51.4.4%
Skull	88.8%	97.3%

The recorded scores on test image Nr. 1. with the latest version.

	Calcifications	Cerebrospinal fluid	Brain matter	Skull
Calcifications	99.8%	0%	0%	0,221%
Cerebrospinal fluid	0 %	25.4%	68.2%	6.33%
Brain matter	0%	0.927%	99.1%	0%
Skull	0%	0%	16.7%	83.3%

	Precision	Recall
Calcifications	99.8%	100%
Cerebrospinal fluid	25.4%	96.5%
Brain matter	99.1%	53.8%
Skull	83.3%	92.7%

The recorded scores on test image Nr. 3. with the latest version.

As it easy to see, the performance of the originally proposed algorithm has the best score on test image Nr. 4, and in the same time the lowest score reached on test image Nr.1. That shows the total percentage of cerebrospinal liquid area and the reached score are in a linear harmonic relation.

Chapter 6

Conclusion

.

5.1 Conclusion

In the present paper - an overlapping clustering algorithm – namely: Fuzzy c-mean was implemented with the aim at segmenting CT scans of the human head.

As a part product of the project a segmentation toolkit and a CT quality improver is implemented.

The experimental result shows that the originally proposed version of FCM is not efficient in distinguishing cerebrospinal liquid from brain liquid.

Even with the extended Population-Diameter Independent version the results are unusable.

After adding more features – based on histogram moments and gray level co-occurrence matrix - the reached scores are increased, but still not satisfying.

As a conclusion, we can say Fuzzy c-mean is not a useful approach of segmenting CT head images, because even with the Population-Diameter Independent version cerebrospinal fluid is misclassified. The reason remains in the percentage of the pixels – representing the cerebrospinal fluid – which are significantly less than the rest of the segmented materials.

5.2 Further works

Some trouble arisen during the project work which delayed the careful testing of the system and the parameter optimalization.

Possible extension of the presented work could use more features. It could be also interesting in an initial step, to segment the image into more clusters and merge the similar cluster into each other. This way the result may be more accurate.

6 Appendix:

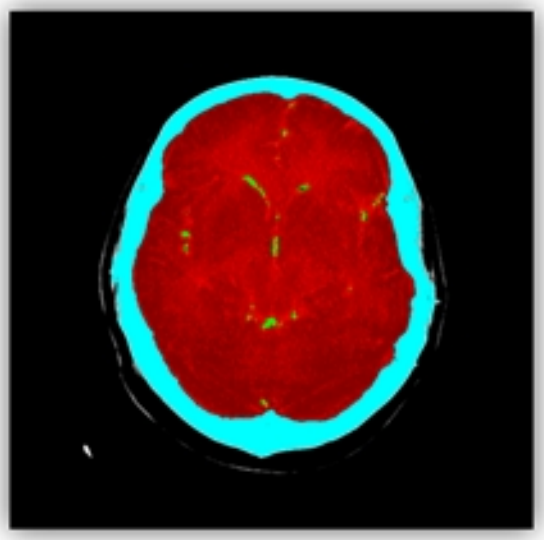


Figure 6.1 Test image No. 1.

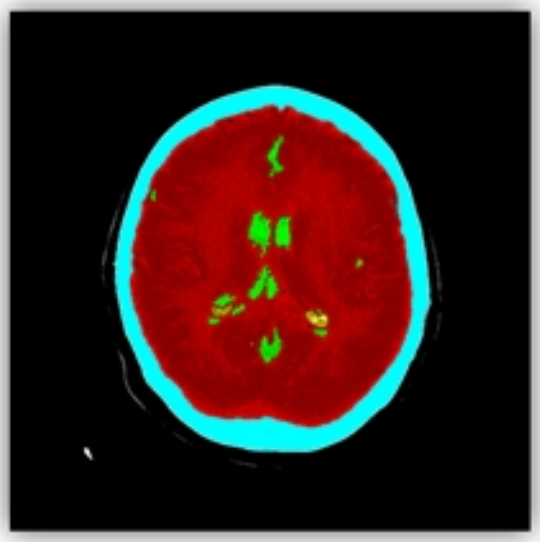


Figure 6.2 Test image No. 2.

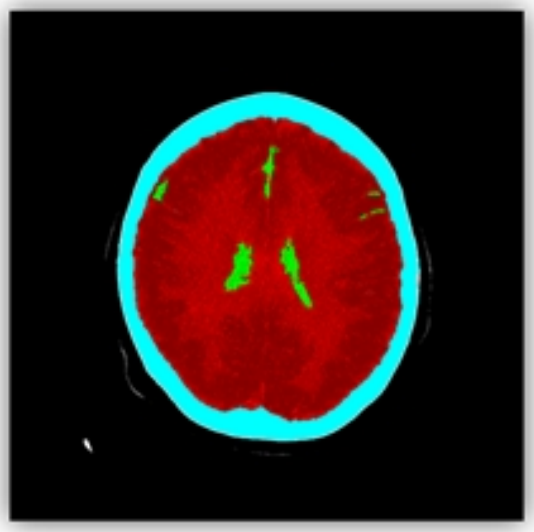


Figure 6.3 Test image No. 3.

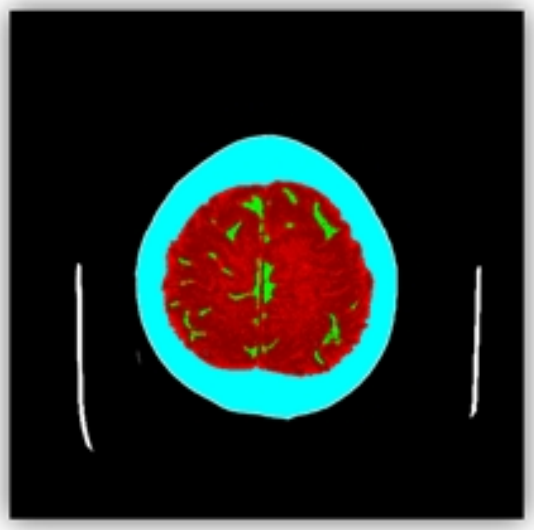


Figure 6.4 Test image No. 4.

Manually segmented images

7 References:

[1.] **Linda G. Shapiro and George C. Stockman** (2001): *Computer Vision*, pp 279-325, New Jersey, Prentice-Hall, ISBN 0-13-030796-3

[2.] **G. Bertrand** (2005): "On topological watersheds", *Journal of Mathematical Imaging and Vision*, Vol. 22, No. 2-3, pp. 217-230.

[3.] **Shihab** (2000): *Fuzzy Clustering Algorithms and Their Application to Medical Image Analysis*. In Ph.D. thesis, University of London

[4.] http://www.connieore.com/2007/09/

[5.] http://www.ams.org/featurecolumn/archive/brain.html

[6.] http://upload.wikimedia.org/wikipedia/commons/3/3b/MRI_brain.jpg

[7.] http://www.ob-ultrasound.net/sliceabd.html

[8.] http://www.answers.com/topic/biomedical-ultrasonics

[9.] http://upload.wikimedia.org/wikipedia/commons/d/d0/Mammogram_showing_breast_cancer.jpg

[10.] http://upload.wikimedia.org/wikipedia/commons/b/bf/Thorax_pa_peripheres_Bronchialcarcinom_li_OF_markiert.jpg

[11.] http://www.ams.org/featurecolumn/archive/brain.html

[12.] http://www.sci.u-szeged.hu/foldtan/CT_SPCEKOLL/CT_alap.pdf