

NEW UNIVERSITY OF LISBON
Faculty of Science and Technology
Integrated Masters Degree in Electrical Engineering and Computers

SEGMENTATION OF STRIATAL BRAIN STRUCTURES FROM HIGH RESOLUTION PET IMAGES

by
Ricardo Jorge Pires Correia Farinha

Dissertation presented at the Faculty of Science and Technology of the New University of Lisbon in fulfillment of the requirements for the Masters degree in Electrical Engineering and Computers

Supervisors: Prof. José Manuel Fonseca
Prof. Ulla Ruotsalainen

Lisbon
2008

Preface

This thesis was carried out during the 2007-08 academic year in the Methods and Models for Biological Signals and Images (M²oBSI) research group, at the Institute of Signal Processing, Tampere University of Technology, Finland.

First, I want to express my deepest gratitude to my parents and family for all the support and encouraging given during my studies. Without them none of this would be possible.

I want to thank my supervisors, Professor Ulla Ruotsalainen (TUT), Professor José Fonseca (FCT-UNL) and senior researcher Jussi Tohka, PhD (TUT) for this opportunity, for their guidance, and for all the support given to me during the elaboration of this thesis.

I thank senior researcher André Ribeiro, PhD. for the patience to read through the thesis and pointing out many useful improvements.

I thank to all my colleagues in M²oBSI research group for always finding time to help and advise me when needed. I specially thank to Antonietta Pepe and Uygur Tuna, my office colleagues, for all the help and friendship.

I wish to thank to Marika Kovanen, in particular, and all my friends for their endless help and encouragement whenever needed.

Lastly, I would like to thank the collaborators in Turku PET Center, University of Turku/Turku University Central Hospital for the outstanding collaboration, for providing useful information on the physical principles behind HRRT scanners and for providing the PET data.

Tampere, May 2008.

Ricardo Farinha

Kärkikuja 2C 76
33720 Tampere, FINLAND

+358 44 94 54 397

Ricardo.pirescorreiafarinha@tut.fi

Abstract

We propose and evaluate fully automatic segmentation methods for the extraction of striatal brain surfaces (caudate, putamen, ventral striatum and white matter), from high resolution positron emission tomography (PET) images. In the preprocessing steps, both the right and the left striata were segmented from the high resolution PET images. This segmentation was achieved by delineating the brain surface, finding the plane that maximizes the reflective symmetry of the brain (mid-sagittal plane) and, finally, extracting the right and left striata from both hemisphere images. The delineation of the brain surface and the extraction of the striata were achieved using the DSM-OS (Surface Minimization – Outer Surface) algorithm. The segmentation of striatal brain surfaces from the striatal images can be separated into two sub-processes: the construction of a graph (named “voxel affinity matrix”) and the graph clustering. The voxel affinity matrix was built using a set of image features that accurately informs the clustering method on the relationship between image voxels. The features defining the similarity of pairwise voxels were spatial connectivity, intensity values, and Euclidean distances. The clustering process is treated as a graph partition problem using two methods, a spectral (multiway normalized cuts) and a non-spectral (weighted kernel k-means). The normalized cuts algorithm relies on the computation of the graph eigenvalues to partition the graph into connected regions. However, this method fails when applied to high resolution PET images due to the high computational requirements arising from the image size. On the other hand, the weighted kernel k-means classifies iteratively, with the aid of the image features, a given data set into a predefined number of clusters. The weighted kernel k-means and the normalized cuts algorithm are mathematically similar. After finding the optimal initial parameters for the weighted kernel k-means for this type of images, no further tuning is necessary for subsequent images. Our results showed that the putamen and ventral striatum were accurately segmented, while the caudate and white matter appeared to be merged in the same cluster. The putamen was divided in anterior and posterior areas. All the experiments resulted in the same type of segmentation, validating the reproducibility of our results.

Table of contents

1. Introduction	6
2. Human striatal neuroanatomy	11
2.1. Striatum	11
2.2. Dorsal striatum	12
2.2.1. Caudate nucleus	12
2.2.2. Putamen	12
2.3. Ventral striatum.....	13
3. Positron Emission Tomography (PET)	15
3.1. Tracer technology	16
3.2. Distribution of the intensity values in a PET image	17
3.3. ECAT HRRT PET scanner	18
4. Deformable models	19
4.1. Medical Image Analysis using Deformable Models	20
4.2. Deformable Curves.....	21
4.3. Deformable Surfaces	23
4.3.1. DSM (Dual Surface Minimization) algorithm	23
4.3.2. DSM-OS (Dual Surface Minimization-Outer Surface) algorithm	25
4.4. Incorporating “A priori” knowledge	26
5. Graph-Based Image Segmentation	27
5.1. Spectral methods for clustering	28
5.2. Normalized cuts	28
5.3. Non – spectral methods	31
5.3.1. Weighted Kernel k-means	31

6. Segmentation of striatal structures	34
6.1. Pre-processing steps	34
6.2. Segmentation of the striatum into four structures	36
6.2.1. Affinity matrix construction methods	37
6.2.2. Clustering methods	39
6.3. Implementation	41
7. Experiments	42
7.1. Material	42
7.2. Results	43
7.2.1. Phantoms	43
7.2.2. Real HRRT images	47
8. Discussion and conclusion	51
References	53
Appendix 1: Eigenvectors and Eigenvalues	59
Appendix 2: Segmentation results	60
Appendix 3: Putamen subdivision	64

Abbreviations

AP – Anterior putamen
BGO – Bismuth germinate
BP – Binding potential
CD – Caudate
CT – Computer tomography
DOI – Depth-of-interaction
DSM – Dual surface minimization
DSM-OS – Dual surface minimization - outer surface
FCT-UNL – Faculty of Science and Technology-New University of Lisbon (Portugal)
FDG – Fluorodeoxyglucose
FMRI – Functional magnetic resonance imaging
FOV – Field of view
H – Hydrogen
HRRT – High resolution research tomograph
IC – Internal capsule
LSO – Lutetium orthosilicate
LV – Left ventricle
MRI – Magnetic resonance imaging
N – Nitrogen
NCUT – Normalized cut
O – Oxygen
OP-OSEM-3D – Ordinary poisson- ordered subsets expectation maximization in full 3D
PET – Positron emission tomography
PP – Posterior putamen
ROI – Region of interest
TUT – Tampere University of Technology (Finland)
2D – Two dimensional
3D – Three dimensional

1. Introduction

The development and proliferation of medical imaging technologies is revolutionizing medicine [1]. Medical imaging can extract clinical information on anatomic structures through computed tomography (CT), magnetic resonance imaging (MRI). Medical imaging also allows obtaining functional information using positron emission tomography (PET), and others, with almost no harmful effects to the body and noninvasively. Besides visualization of anatomic structures, medical imaging is now also used to do surgical planning and simulation, intra-operative navigation, radiotherapy planning, and to track the progress of diseases [1]. E.g., radiotherapy can subject a tumor to a specific dose of radiation with minimal collateral damage to healthy tissue.

Each imaging modality was developed to provide information on particular aspects of the imaged object. X-ray based computed tomography (CT) produces images based on the photon attenuation as the X-rays goes through the tissue. Magnetic resonance imaging (MRI) imaging measures the proton or water density. MRI and CT images are used to visualize the anatomical structure of the subject, since they allow distinguishing between different tissues. Positron emission tomography, presented in more detail in chapter 3, and functional MRI imaging (fMRI) inform on the functional properties of the tissue. In figure 1.1 are shown cross-sections acquired with MRI and PET.

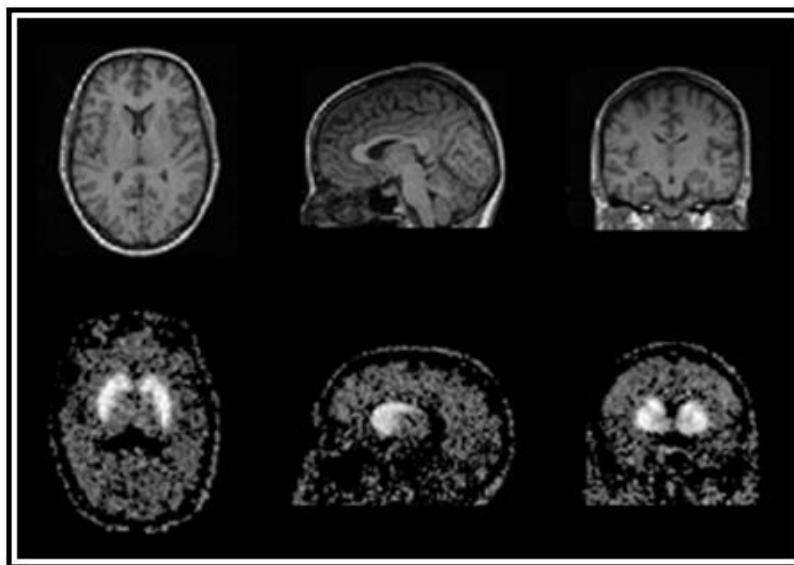


Figure 1.1. Examples of medical images from the head, transaxial, sagittal and coronal views. (top) A T1-weighted MRI image, (bottom) PET image. The images are from different subjects. The images were provided by Turku PET Center.

Although modern imaging devices attain high quality anatomical views, the analysis of the image data via computational methods is still very limited, and requires human intervention. The development of more advanced computational methods will allow extracting and interpreting quantitative data efficiently and in a repeatable way, which can then be used to support a wide range of tasks, from diagnosis to clinical interventions.

One aspect in which computational methods can contribute is in image segmentation. For this to be possible, it is necessary to accurately determine the boundaries of each structure. This has been shown to be unattainable by traditional low-level image processing techniques since they only account for local information, giving rise to incorrect assumptions during the integration process, which leads to generating unrealistic object boundaries.

Most clinical segmentation is currently done by manual slice editing, that is, a human operator manually delineates the region of interest on each 2D slice of the 3D image. This process has several disadvantages, for example, the results will be dependent on any bias of the operator, when viewing each 2D slice and deducing the spatial shape of 3D structures. Also relevant to the results is the fatigue of the operator, due to the enormous amount of data that must be manually or semi-automatically extracted. This makes the process expensive, if feasible, and, more importantly, the final result might not be easily reproducible by another operator, hampering the comparability between the segmentation results.

Automation of medical image analysis is a highly complex, not only due to the complexity of the shapes one wants to extract but also due to the variability within and across individuals. The noise in medical images results from a complex combination of noise from several sources, e.g. measurement noise, natural intensity variation within structures of interest, and modality dependent imaging artefacts [2]. This makes the task of suppressing noise and artificially correcting its effects a difficult task. Since different imaging modalities use different measurement devices, they are subject to different sources of noise and differ in noise levels. Importantly, the effects of noise can vary in different structures of interest which causes additional difficulties in their discrimination, since discrimination is done calculating the differences in intensity values rather than intensity values themselves [2].

The aim of this thesis is to improve and extend the methods presented in [3]. There, the authors proposed an automatic method to extract striatal structures, namely, caudate and putamen, from binding potential PET images. The data acquisition was done using the (ECAT EXACT HR+, Siemens CTI, Knoxville, TN) scanner and on the dedicated GE Advance PET scanner (GE Medical Systems, Waukesha, WI). The methodology in [3] can be divided in the following steps:

1. Extract the brain surface using a deformable model.
2. Extract hemisphere images from the extracted brain surface.

3. Extract the left and right striatum from the left and right hemisphere images using a deformable model.
4. Extract putamen and caudate from both right and left striata.

The deformable models have been used to segment, visualize, track, and quantify a variety of anatomical structures (see chapter 4 for a more detailed description) and a general overview on deformable models. The methods used to extract the information from the images in steps one to three are explained in chapters 4 and 6. The extraction of the putamen and the caudate in step four is done using a 2-way normalized cut algorithm [3]. In chapter 5 is presented an overview on the Graph-Based image segmentation, and more specifically, on the normalized cuts algorithm.

The development of the new brain-dedicated, high-resolution PET scanner (ECAT HRRT, Siemens Medical Solutions, Knoxville, TN) that provides images with enhanced resolution, step four in [3] fails to segment the striatal structures. The lack of success was due to the excessive computational complexity of the segmentation algorithm. This excessive computational complexity became prohibitive due to the enhancement of the resolution of the new HRRT scanner. This resolution enhancement allows distinguishing more than two structures in the PET data. Therefore, the method used in [3] to extract the striatal structures (2-way normalized cut algorithm) became obsolete since it can only manage the separation into two structures. In this study, we aim to overcome this limitation and develop an algorithm capable of separating the striatum into four structures (caudate, putamen, ventral striatum, and white matter). Due to the limited success of the previous method, this thesis proposes the use of a mathematically identical method, but that is able to dwell with the increased complexity of the new type of high resolution images.

In short, this thesis presents new automatic techniques to successfully extract striatal brain structures, namely putamen, caudate, and ventral striatum, from PET HRRT binding potential images. In chapter 2, information is given about the structures that the proposed method aims to extract. The automatic method consists of a series of consecutive steps shown in the hierarchy graph below (figure 1.2).

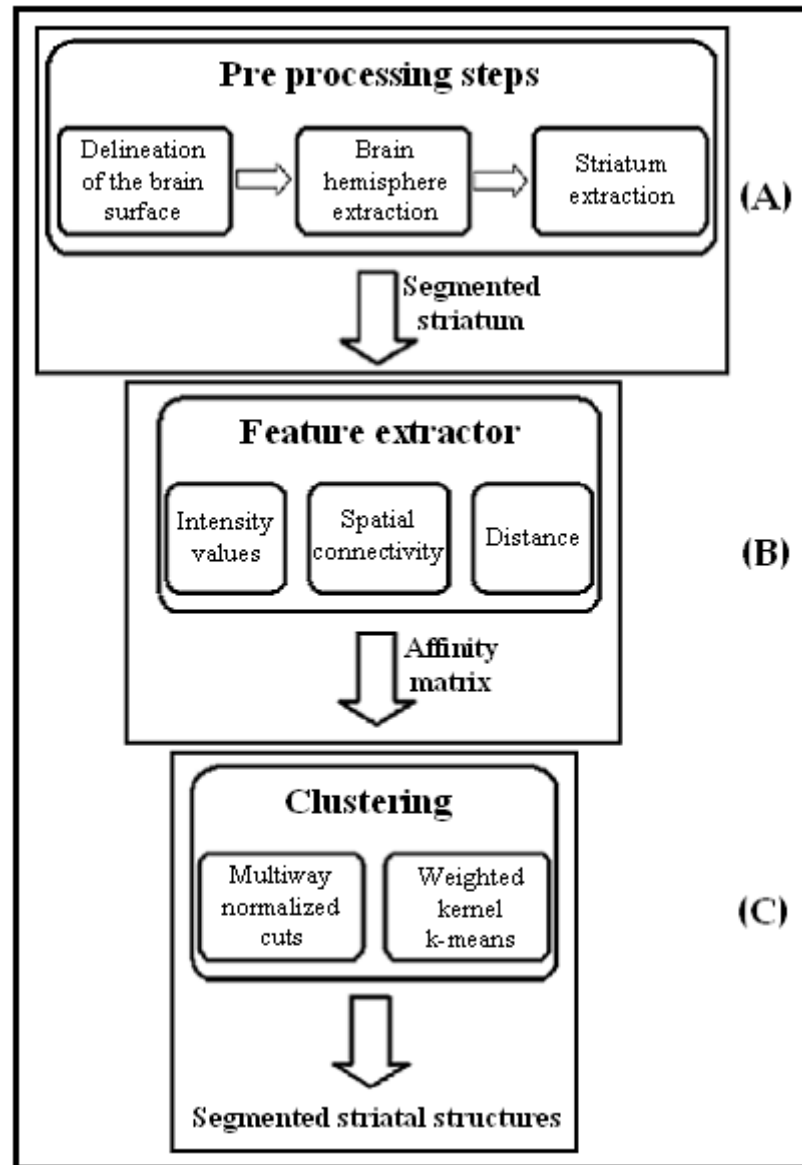


Figure 1.2 Steps of the segmentation of caudate, putamen, ventral striatum. (A) Preprocessing steps, (B) affinity matrix construction method, (C) clustering process.

The preprocessing step (figure 1.2 (A)) consists on the extraction of the striatum, from HRRT PET binding potential images. This extraction method is the same as the one used in [3] and is described above. However, some tuning of the initial parameters was necessary.

Once the striatum was extracted, several features of the striatal voxels are computed (figure 1.2 (B)). This information is then stored in an affinity matrix and provided to the clustering algorithm. The image features used are Euclidean distances between voxels, spatial connectivity and intensity values. The affinity matrix contains the similarity values between all pairs of voxels. That is, in position i, j is written the similarity value between voxels i and j , thus, this

matrix is symmetric and all values in the diagonal are equal to 1, see section “affinity matrix construction methods” for the methodology and an example. Determining the affinity matrix is a crucial step to guarantee enough smoothness so that the clustering process can provide a good segmentation.

Finally, the clustering algorithm segments the image (figure 1.2 (C)). The clustering procedure is treated in this thesis as a graph partition problem. Two approaches were developed that are capable of partitioning the graph: the multiway normalized cuts and the weighted kernel k-means. In both methods, the affinity matrix represents the graph that one wants to partition into different regions. The strategy used in the multiway normalized cut algorithm is to apply the k-means clustering algorithm to the computed graph eigenvalues in order to partition the image into connected regions. On the other hand, the weighted kernel k-means classifies, iteratively, each voxel as belonging to a cluster. This classification is done using the information contained in the affinity matrix (graph) of the image. Although this procedure allows successively extracting the desired structures of interest, it has some subtleties. Namely, a careful initialization is required in each step, to obtain maximum accuracy, i.e., the final result depends on the number of desired clusters, initial estimative for the clustering result, maximum distances and spatial connectivity allowed between voxels and, number of iterations. These parameters have to be carefully tuned to provide the most accurate result. The same values were found to be the optimal for all tested HRRT PET image, thus no further tuning was required. Chapters 6 and 7 give, respectively, a detailed description of the methodology and results. In chapter 8, the discussion and conclusions of the results are presented.

2. Human striatal neuroanatomy

The aim of this thesis is to develop efficient automated methods to extract striatal brain structures from high resolution PET images. To this aim, in this chapter, the structures of interest are introduced.

The brain is the center of the human nervous system and the main regulator of the ongoing processes in the human body. The human brain is composed of white and gray matter and is surrounded by three layers of special soft tissue. The white matter is the area of the brain whose nerve fibers are involved in myelin responsible for the coloration [4]. The gray matter consists mainly of nerve cell bodies (neurons). Its gray brown color is due to the neuronal capillary blood vessels and the neuronal cell bodies [4].

The human brain is centrally divided into two hemispheres that are connected by bridges of white matter. The brain structure is usually characterized by dividing it six lobes, depending on the function to which they have been associated to. Four lobes are visible from the outside, namely, the frontal, the parietal, the occipital and, the temporal lobe, while the other two are internal: the insula and limbic lobes. The frontal lobe is mainly associated with reasoning, planning, parts of speech, movement, emotions, and problem solving. The parietal lobe is associated with movement, orientation, recognition, perception of stimuli. The occipital lobe is associated to visual processing and the temporal lobe is associated to perception and recognition of auditory stimuli, memory, and speech [5].

2.1. Striatum

The striatum is an individual entity with a toric topology. There is one striatum in each brain hemisphere. The striatum is composed by several structures (figure 2.2): the caudate, the putamen and the ventral striatum. The caudate is the closest to the center of the brain, the putamen the furthest away and the ventral striatum is beneath them. The putamen and caudate are separated by the anterior limb of the internal capsule [5], however there are some bridges of white matter connecting these structures. These structures are all contiguous, making difficult the delimitation of the ventral striatum. A debate is still ongoing within the medical community to delimitate its boundaries.

The striatum is best known for its role in the planning and modulation of movement pathways but is also involved in a variety of other cognitive processes involving executive functions. In humans, the striatum is activated by stimuli associated with reward, but also by aversive, novel, unexpected or intense stimuli [5].

2.2 Dorsal striatum

2.2.1. Caudate nucleus

The caudate nucleus has an arched geometry that can be separated into two main structures, namely, the head and the tail also referred to as the body of the caudate. The head is located in the anterior portion of the caudate and it is an enlarged area when compared to the tail. The head of the caudate nucleus and the putamen are horizontally separated with a tract of white matter and, the bottom of each of these structures is vertically connected by the nucleus accumbens. The tail is connected to the head of the caudate and it elongates circularly, passing through the back of the brain. The tail forms the roof of the inferior horn of the lateral ventricle.

Recently, it was demonstrated that the caudate is highly involved in learning and memory, particularly regarding feedback processing [6]. It is known that neural activity occurs, in general, in the caudate while an individual is receiving feedback. The left caudate has been suggested to be related with the thalamus, which governs the comprehension and articulation of words as they are translated between different languages [7].

2.2.2. Putamen

The putamen is the largest striatal structure. This structure is cytologically uniform, but is not uniform in terms of the functional behavior of its components. The putamen complex is divided into discrete striosomes, embedded in a background matrix, the extrastriosomal matrix [5]. The matrix and striosome have several chemical differences, e.g. if the putamen is stained immunocytochemically with enkephalin one can observe that this compound will be concentrated in the areas where the striosomes are, which is around the striosomes periphery [8]. The putamen has various types of connections and neurotransmitters. It receives most of its inputs from the motor and somatosensory areas of the cortex. The putamen outputs are delivered mainly to the globus pallidus and thalamus. This information will then be delivered to the motor, premotor and supplementary motor areas. Therefore, the putamen is considered to be a major player in most of the striatal motor functions [5].

The putamen can be divided into two sub regions, the anterior and posterior putamen (figure 2.1), as defined respectively as ventral and dorsal putamen [9]. The anterior putamen is composed by gray matter which is bounded by the anterior and posterior limbs of the internal capsule, and, laterally, is bounded by the external capsule in the first three planes ventral to the plane containing the dorsum of the putamen. The posterior putamen is bounded by the pallidum, by an external capsule, and by the internal capsule [9].

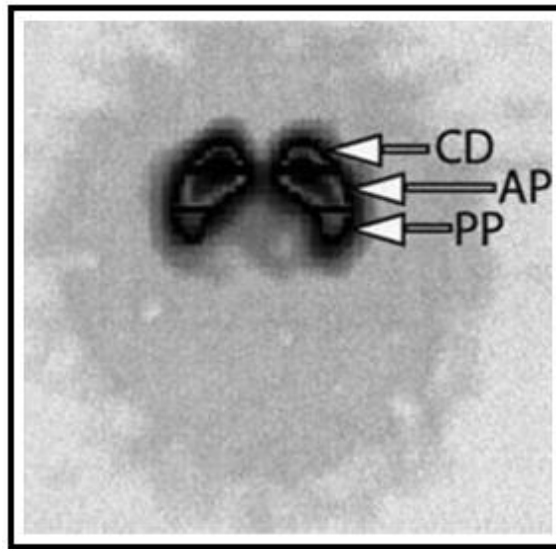


Figure 2.1 Striatum subdivisions, (CD) caudate, (AP) anterior putamen, (PP) posterior putamen. The image is from [10].

2. 3. Ventral striatum

In the literature, the striatum is usually divided into two areas, the ventral striatum and the dorsal striatum [11]. The dorsal striatum is the sensorimotor-related part of the striatum and it is composed by the putamen-caudate complex. The ventral striatum is related with emotional and motivational aspects of behavior [12]. Dysfunctions of the ventral striatum have been associated among individuals with schizophrenia [13], obsessive-compulsive disorder [13], depression, and drug addiction [14].

So far, no structural or functional marker has been identified that allows an accurate distinction between ventral and dorsal striatum. Commonly, this separation is assumed to be between the caudate-putamen complex and the core and shell structures of the nucleus accumbens (figure 2.2). However, the ventral striatum is not composed only by the nucleus accumbens, but it extends ventrally through the caudate and putamen [11], [15].

The nucleus accumbens is divided into two areas (shell and core structures) according to their specific function [16]. The shell is the structure located at the lower bottom of the nucleus accumbens (figure 2.2). The shell is known to be involved in the expression of certain innate, unconditioned behaviors, such as feeding and defensive behavior [17]. The core is located above the shell and it is responsible for response-reinforcement learning. The shell and core subregions play important but distinct roles in instrumental conditioned learning that may be potentiated with psychostimulants [16].

In short, the ventral striatum, a specific region of the striatum, is the structure responsible for connecting the limbic, cognitive, and motor systems within the striatum. Experimental studies in animals have shown that the ventral striatum plays an important role in several forms of behavioral learning [18].

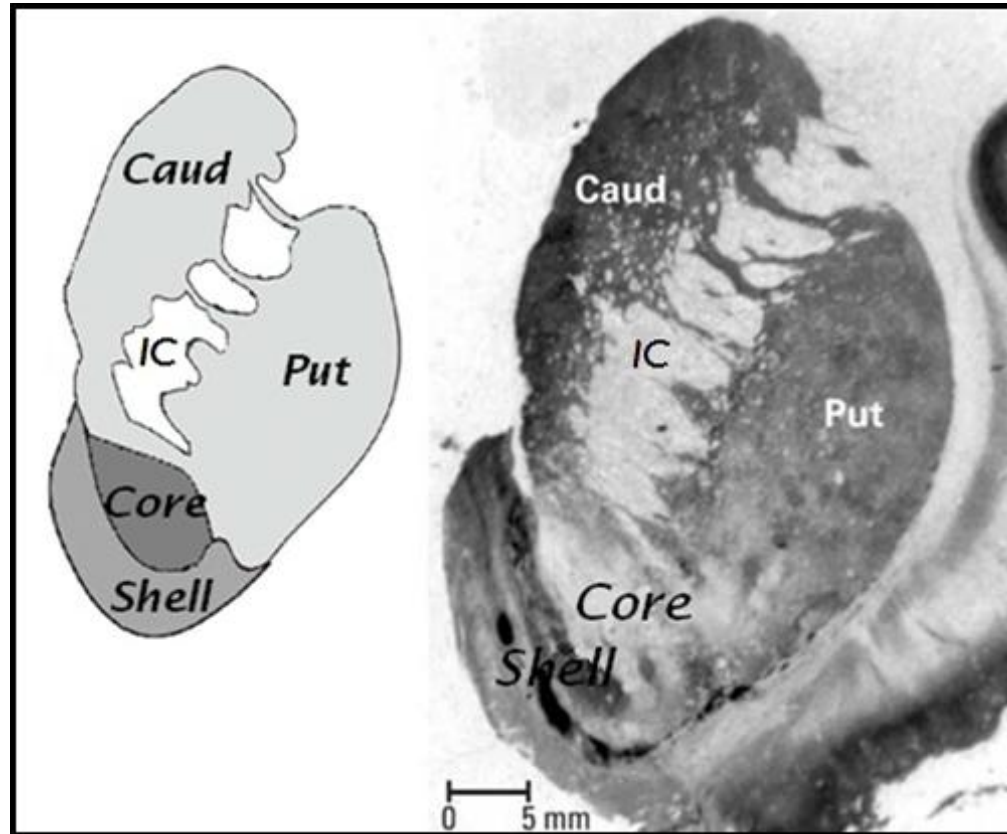


Figure 2.2 Right striatum, Core and Shell belonging to the nucleus accumbens we can consider the nucleus accumbens as the ventral striatum, Put (putamen) and Caud (caudate), IC (internal capsule). The image is from [18].

3. Positron Emission Tomography (PET)

Positron Emission Tomography (PET) is the medical imaging technology used to acquire the images to which the proposed methods are applied. In this chapter a theoretical introduction to its principles is given.

Positron emission tomography (PET) is a widely used medical imaging technology. Within the several categories of medical images, PET is categorized as a functional imaging modality. This means that PET imaging devices have the ability to observe physiological functions in vivo, and from these create three dimensional dynamic models. This functional information provides the enhancement of the understanding that the medical community has of the biochemical basis of normal and abnormal bodily functions [19]. PET can be used to explore the functioning of the internal organs of both animals and humans.

PET imaging integrates two technologies: the tracer kinetic assay method and computed tomography (CT). The tracer kinetic assay method consists of introducing a radio labeled biologically active compound (tracer) into the subject blood flow, and in using a mathematical model to describe the kinetics of the tracer as it participates in a biological process. The tracer kinetic model requires information about the tissue tracer concentration, measured by the PET scanner [20].

During PET acquisition, one administrates into the patient, by injection or inhalation, a positron-emitting compound containing a radioactive isotope. This compound will concentrate in specific regions of the body, with a given (unknown) density. When one emitted positron encounters an electron, they annihilate each other and produce two photons of 511 keV. The photons will move in opposite (180°) directions, to conserve momentum. The energy of these photons is usually sufficient to allow the photons to “escape” the subject body, being then detected by the detector rings that surround the subject (figure 3.1 (B)).

For an emission event to be counted, both photons must be registered nearly simultaneously (within a time window τ) at two opposite detectors (scintillation crystals composed of e.g. bismuth germanate (BGO)), i.e. they must occur on the same coincidence line (figure 3.1 (A)) [19]. The absolute value of the counts in a PET scan results from the counts of true coincidences plus the counts of random coincidences. True coincidences are coincident events that originate from a single annihilation process, while random coincidence events occur when, accidentally, a pair of detectors located in opposite positions record an uncorrelated signal within a time window τ .

The detectors of the PET scanner are arranged in cylindrical shaped rings. This property allows the scanner to register the events in three dimensions. From this set of data, the spatial

distribution of the radioactivity in the body can be reconstructed by using appropriate reconstruction algorithms [21]. All measured coincidence events (valid events detected by a pairs of detectors) are stored into sinograms. A sinogram is a two-dimensional matrix where the columns store each image projection angle and the rows store the spatial positions within the (measured) projections, i.e. they are projection views at different angles [22]. The reconstruction algorithms take into account all the coincidence events stored in the sinogram, at all angular and linear positions, when reconstructing the image.

The image quality in PET is strictly related with the applied image reconstruction method. The positron emission data acquisition is not free from noise, namely, there is a substantial amount of statistical noise that arises from the statistical nature of the decay of the positron emitted by radio-labeled isotope [23].

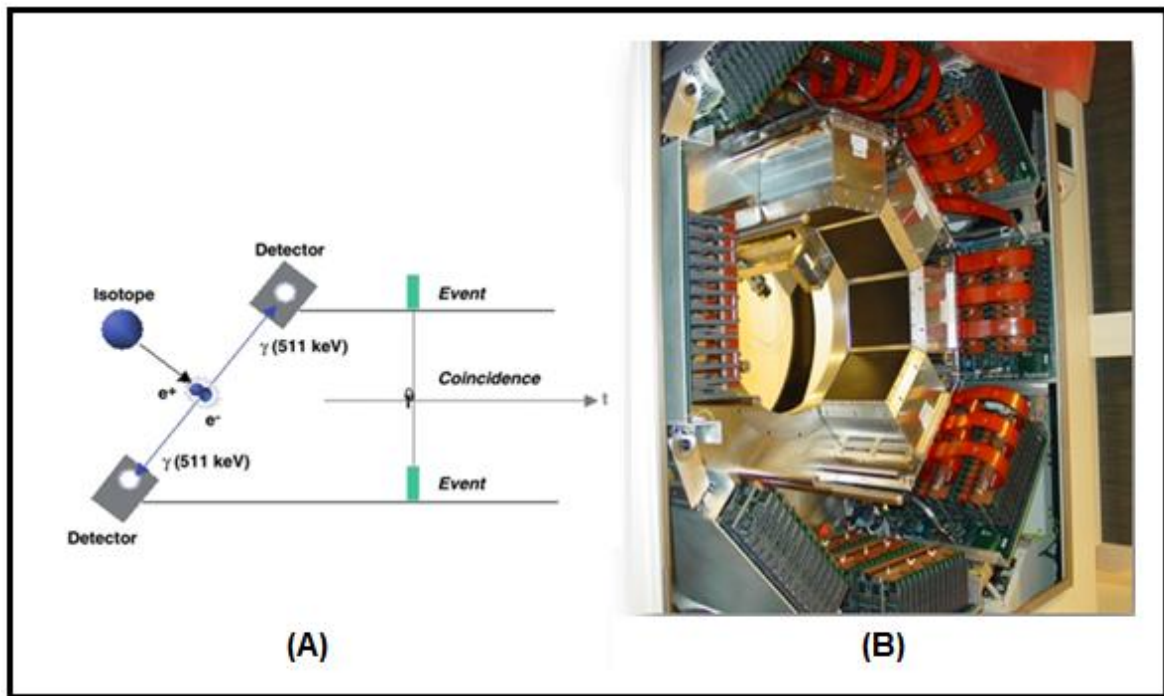


Figure 3.1. (A) [^{18}F] annihilation and the positron emission tomography (PET) coincidence principle, (B) interior of the HRRT scanner chamber. The images are from [21] and (from presentations of J. Johansson).

3.1. Tracer technology

The tracer technology is one of the fundamental keys in PET imaging. The main idea behind this technology is that before the compound is injected into the subject's blood flow, an artificial isotope has to replace an atom of the compound. This isotope is a radio-labeled tracer artificially engineered that, when injected in the blood flow, can be detected from outside the body. The

tracer does not change the chemical properties of the compound, ensuring that the biochemical process to be studied is not altered.

The atoms that are usually artificially replaced by their isotopic forms are carbon (C), hydrogen (H), nitrogen (N), and/or oxygen (O). These stable molecules are replaced respectively by ^{11}C , ^{13}N or ^{15}O . These radiolabeled isotopes, whose decay results in the emission of a positron, are ideal for this purpose due to their short half-life, respectively (^{11}C half-life is 20 min; ^{13}N , is 10 min; ^{15}O , is 2 min) [21]. However, their short life expectancy requires their production to be quickly followed by their introduction in the patient, a difficult and rigorous task. For this reason, the production of these compounds by the cyclotron is usually performed near the place of application into the patient.

Raclopride labeled with ^{11}C is used as a PET tracer to study the function of dopaminergic synapses. Raclopride binds to dopamine D2 receptors and is a selective, reversible inhibitor of dopaminergic D2 receptor function [20].

The radiation generated by the injected compound is absorbed by the subject tissue according to its mass. This absorption can be quantified by the PET scanner or by computer tomography (CT). The resulting dynamic data can be collected from specific biological structures (regions of interest) by recording its radioactivity over time [21]. This recording is done in predefined time frames in which the analyzed targets are scanned.

To analyze the results from the scans, one needs an estimation of how the tracer binds to the region of interest receptors. This information is stored in the so called parametric binding potential (BP) images. By definition, the binding potential is a measure of the ratio of concentration of specific binding and is a reference concentration at equilibrium [24]. These images are obtained from the recorded dynamic data, through the use of a predefined model, e.g., the “simplified reference tissue model” [25].

3.2. Distribution of the intensity values in a PET image

The spatial distribution of intensity values within an image depends on the radiopharmaceutical applied [23], that is, on the uptake areas defined by the tracer. The uptake area of a receptor-like tracers is the one with the highest concentration of the tracer type receptors. For example, if the tracer used is [^{11}C] raclopride, the higher uptakes will be found in the striatal area because it is the region with the highest densities of dopamine D2 receptors in the human brain, and therefore it will have the highest intensity values. In short, image volumes with higher intensity values than its surroundings are classified as the functional region for the applied tracer [23].

3.3. ECAT HRRT PET scanner

The ECAT HRRT (High Resolution Research Tomograph) is a dedicated 3D-only brain scanner. This tomography device is imbued with the new LSO scintillators. It uses depth of interaction (DOI) information to achieve uniform isotropic resolution across a 20cm diameter volume [26].

In Positron Emission Tomography there is a continuous effort to increase the sensitivity and to improve spatial resolution up to its physical limits. For high sensitivity a compact geometry is needed. Also good spatial resolution requires a small diameter of the detector arrangement to reduce the effect of the angular deviation from 180 degrees of the two annihilation photons. A compact geometry entails an increase in scatter and random coincidences. Random rates can be reduced by a shorter coincidence window, which requires a fast crystal. The other prerequisite for improved resolution is a reduction in crystal size, which necessitates a crystal material with high luminosity. If good resolution shall be preserved over a large volume in a scanner with a compact geometry and a small crystal size, depth of interaction (DOI) needs to be measured additionally.

The High Resolution Research Tomograph scanner was designed to meet all these challenges. It employs the new scintillator LSO (Lutetium Oxy Orthosilicate). Its high light yield allows a reduction in crystal size and its short scintillation rise and decay time a short coincidence time window and a reduction of the dead time.

The single crystal layer in HRRT scanner consists of eight flat panel detector heads, arranged in an octagon. The distance between two opposing heads equals 46,9cm. Each head contains 9x13 LSO crystal blocks of 7.5mm thickness, which are cut into 8x8 crystals, resulting in 7488 individual crystal elements per head and 59 904 crystals for the entire gantry. The size of one crystal equals 2.1x2.1x7.5mm. The HRRT scanner has a 31.2cm transaxial and 25.5cm axial FOV and a gantry port diameter of 35cm.

In [3], one of the scanners used for the acquisition was the Exat Ecat HR+. This scanner was running in 3D mode during acquisition. From the scan, 13 frames were obtained, with dimensions of 128x128x63 slices and voxel size of $2.11 \times 2.11 \times 2.11 \text{ mm}^3$. In comparison, the HRRT scanner extracted images whose size is 200x200x150 slices with an isotropic voxel dimension of $1.22 \times 1.22 \times 1.22 \text{ mm}^3$, making the automatic analysis of this data a more defying task when using the HRRT scanner due to the size of the images to extract.

4. Deformable models

In this chapter an overview is presented on deformable models. These models are used in the striatum extraction from the high resolution images.

The development of digital imaging technologies made possible the use of automated techniques of imaging processing, which have an increasingly crucial role in medical imaging. Image segmentation is now used at a daily base in medical environment. There are many different applications for image segmentation, e.g. quantification of tissue volumes, diagnosis, localization of pathology, study of anatomical structures, treatment planning, and computer-integrated surgery [27].

There is no standard approach to a segmentation problem. The adequate segmentation procedure varies according to the data and the application in use. One possible definition for image segmentation is: “Image segmentation refers to a process of dividing the image data into regions whose voxels have some common property such as uniform gray-level. A region is a group of voxels with similar properties. Regions often correspond to or have a link to objects” [28].

The characteristics of the data and the inexistence of a method to assess the validity of the result of the segmentation requires the use different segmentation algorithms. As a consequence, new segmentation problems require experimental validation of the segmentation algorithms used for that particular problem. Moreover, image segmentation is not yet an easy task in terms of computational complexity. Although the increasing computation power continuously improves the state of the art, several problems remain whose solution is still very challenging, namely, removal of intrinsic noise of the images, removal of image artifacts, improvements of image resolution, etc. These problems arise due to the object shape and image quality variability and the limited resolution allowed by the scanners. Therefore, when applying classical segmentation techniques, for example, edge detection or “thresholding”, they have a high probability of failure. When using these classical techniques, one has to apply sophisticated image preprocessing or postprocessing techniques to obtain good segmentation results, decreasing the failure rate [27].

The most popular image segmentation methods used for intensity images (image represented by point-wise intensity levels) are: thresholding, edge-based methods, region-based methods, and deformable models. The threshold techniques base their decisions on local voxel information and are suitable to segment objects that fall outside the range of intensity levels of the background. Spatial information is ignored in these methods, decreasing their robustness. Some examples of the application of these methods can be found in [29], [30]. The edge-based methods were developed to detect contours. These methods are “fragile” when the images have broken contour lines. Moreover, their robustness is also highly reduced if there are blurred regions. Some

examples can be found in [31]. Region-based methods rely on grouping neighboring voxels of similar intensity levels to segment the image, thus, the image is partitioned into connected regions e.g. [32]. Deformable models usually start with some given boundary shape formed by spline curves, and iteratively modifies its topology by applying various shrink/expansion operations according to a predefined energy function e.g. [3; 33]. There is a risk that the computation of the energy minimization becomes trapped in some local minimum of the energy function. Fortunately, optimization algorithms can be added that solve this problem.

Deformable models are a successful approach to image analysis, by combining geometry, physics, and approximation theory concepts, and have proven to be effective in segmenting, matching, and tracking anatomic structures, being capable of coping with the diversity of biological structures [1].

Deformable models can be subdivided into two subcategories, deformable curves and deformable surfaces, respectively, to solve two dimensional or three dimensional image segmentation problems. These curves or surfaces are defined inside the image plane or volume, and will be deformed until the desired segmentation result is reached. To obtain the deformations, internal and external forces are applied. The internal forces are defined within the curve or surface and are designed to smooth the model during deformation. The external forces depend on the image data and are applied to move the curve or surface toward an object boundary or other desired features within an image.

By introducing a priori knowledge about the shape of the object of interest, deformable models become more robust to noise and boundary gaps. This initialization makes deformable models very suitable for being applied in medical imaging, because the human anatomy is well known. Deformable models have a low sensitivity to their initialization, e.g. assumed initial surfaces or curves, initial anatomic information, etc. This quality allows using automatic methods without the need of human intervention, thus making them suitable to analyze large sets of data. On essence, this is an iterative method whose goal is, at each step, to find the global minimization of the surface energy function [23].

4.1. Medical Image Analysis using Deformable Models

Deformable models were first developed for computer vision and computer graphics problems. However, their potential when applied to medical image analysis was soon discovered. Deformable models show increased accuracy, precision, reproducibility when compared with the limited manual slice editing and traditional image processing techniques. These models can be used in a variety of medical imaging technologies such as X-ray, computed tomography (CT), angiography, magnetic resonance imaging (MRI), ultrasound and, PET. Deformable models have

been used to segment, visualize, track and quantify a variety of anatomical structures. These include the brain, heart, face, cerebral, coronary and retinal arteries, kidney, lungs, stomach, liver, skull, vertebra, brain tumors, fetus, and even cellular structures such as neurons and chromosomes [1].

Deformable models consider the object boundaries as a whole and can make use of “a priori” information about the object shape to aid in the segmentation problem. The inherent continuity and smoothness of the models can compensate for noise, gaps and other irregularities in object boundaries. Furthermore, the parametric representations of the models provide a compact, analytical description of object shape. These properties make this technique robust and elegant in linking sparse or noisy local image features into a coherent and consistent representation of the object.

4.2. Deformable curves

The first deformable models were used to segment two dimensional structures. One of the first of these deformable models to be used was snakes [34] (figure 4.1). These algorithms are usually initialized by the user. A deformable contour (figure 4.1 (C)) is first defined. Next, this contour expands until it reaches the boundaries of the region of interest (figure 4.1 (d), (e), (f)). The final result of the segmentation procedure can still be tuned by the user. This tuning is done by changing the models parameters until a smoother segmentation result is obtained. The final segmentation consists in a single two dimensional slice that can be used as a first guess in neighbouring image slices, sparing the user the burden of initializing manually the contour of each image slice. If there is a sequence of two dimensional image scans, one can analyze them individually. After segmenting all the image slices, the information accumulated can then be aggregated into a continuous three dimensional surface model.

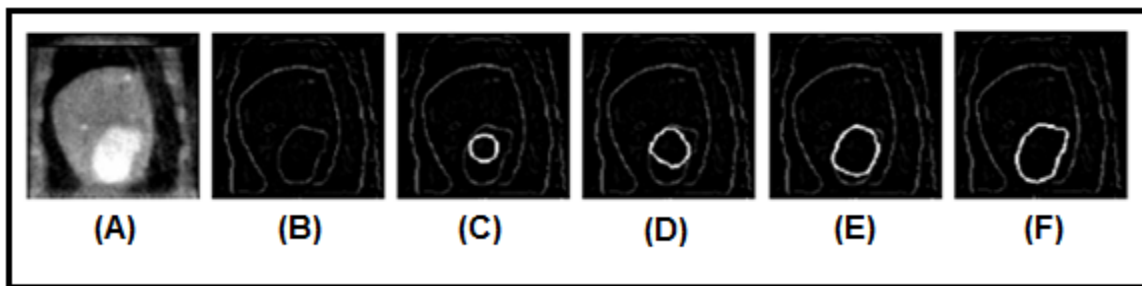


Figure 4.1. (A) Intensity CT image of LV. (B) Edge detected image. (C) Initial snake. (D)-(F) Snake deforming towards LV boundary, driven by “inflation” force . The images are from [33].

However, these two dimensional segmentation approaches have some limitations. As mentioned, they are designed to be interactive methods; however in most medical applications it is necessary to initialize the contours automatically near the regions of interest limiting their applicability. Another limitation is that the objective function, used to allow these methods to expand automatically, can constrain their geometric flexibility and prevent the curves/contours from representing long tube-like shapes or shapes with significant protrusions or bifurcations (figure 4.2) [1]. Classic methods are topology dependent, i.e. they are parametric models and one must provide in advance topological information about the structure of interest.

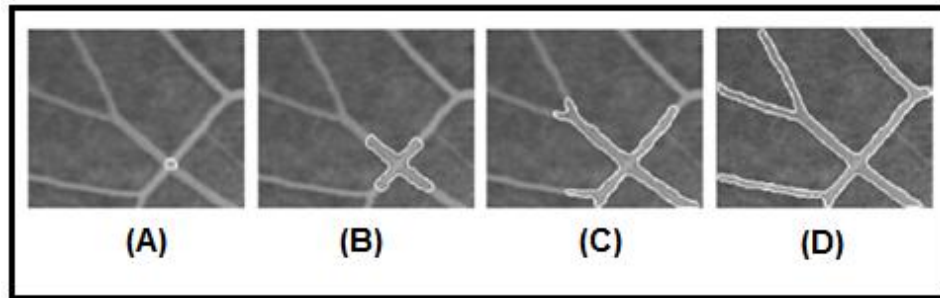


Figure 4.2 Image sequence of clipped angiogram of retina showing automatically subdividing snake flowing and branching along a vessel. The images are from [33].

Due to these limitations, more advanced methods that can automatically cope with topological changes were developed. These shape modelling schemes, topology independent, can be deformable contours or deformable surface models. They can extract long tube-like shapes or shapes with bifurcations (figure 4.2), and also sense changes in the topology (figure 4.3).

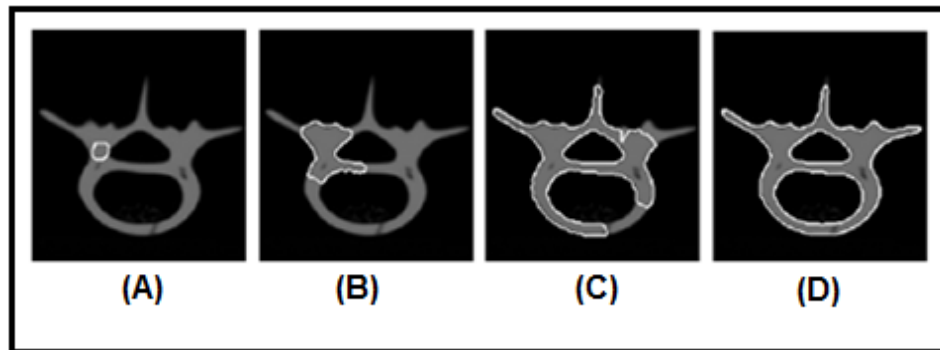


Figure 4.3. Segmentation of a cross sectional image of a human vertebra phantom using a topology-adaptable snake. The snake begins as a single closed curve and becomes three closed curves. The images are from [33].

4.3. Deformable surfaces

Segmenting 3D image volumes slice by slice, both manually or applying 2D contour models, is a laborious process and requires a post-processing step to connect the sequence of 2D contours into a continuous surface [35]. There are several problems associated with the use of this technique. The most important one arises due to the lack of contextual slice-to-slice information, when analyzing sequences of adjacent 2D images. This lack of information results in inconsistencies in the reconstructed surfaces, such as the appearance of bands or rings.

However, the use of fully three dimensional deformable models instead of using a set of individual contours is a faster and more robust segmentation technique, whose results do not suffer from lack of contextual information, which ensures smoothness and coherent surfaces between image slices.

4.3.1. DSM (Dual Surface Minimization)

The DSM (Dual Surface Minimization) algorithm is a deformable surface model. This algorithm is initialized with two meshes, respectively, the inner and outer meshes. A mesh is defined as a volume composed by a set of edges and vertexes. The inner mesh is located inside the region that one wants to extract and the outer surfaces are placed outside the regions of interest.

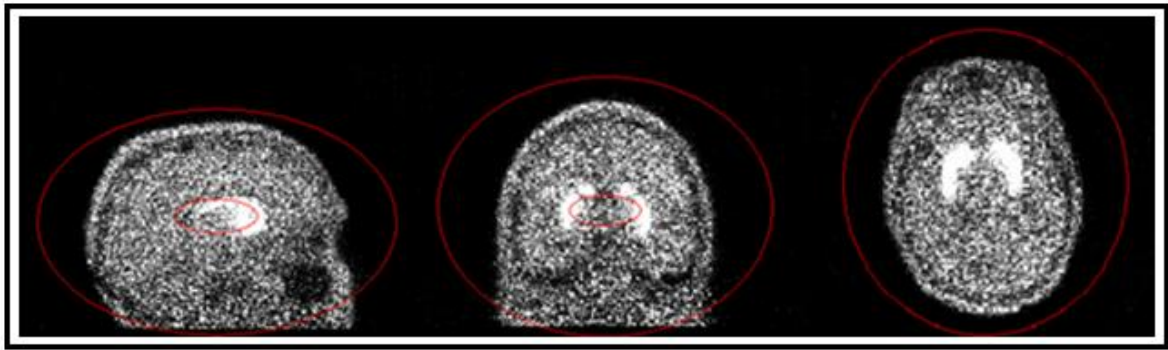


Figure 4.4 HRRT PET Scan, Sagittal, Coronal and Transverse views, image provided by Turku PET Center.

A two dimensional representation of the initial meshes is shown in figure 4.4 as a red outline. The volume that lies inside both initial surfaces (meshes) is called “search volume” and is the area to which the minimization process will be applied to, i.e. where the initial meshes will shrink or expand.

This algorithm is an iterative method that in each iterative step calculates the energy of the current mesh. The algorithm strives to minimize the energy, giving as output the surface with lower energy, thus it is a typical energy minimization problem. The energy term on the DSM algorithm is divided in two terms, the external and the internal energies. The external energy depends on the underlying image, and its minimization is achieved by drawing the smallest surface that still contains all the salient features of the image while the internal energy controls the shape of the surface [23].

The global energy of the surface informs us of how well the mesh approximates the surface of interest and it is defined as:

$$E(M) = \frac{1}{N} \sum_{i=1}^N [\lambda E_{\text{int}}(m_i | m_{i_1}, m_{i_2}, m_{i_3}) + (1 - \lambda) E_{\text{ext}}(m_i)] \quad (1)$$

In (1), E_{int} is the internal energy, E_{ext} the external energy, the parameter λ is a regularization parameter that controls the trade off between the internal and external energies. M is a simplex mesh and its denoted by $M = \{m_1, \dots, m_N\}$, where $m_i \in \mathbb{R}^3$ are coordinates of the meshes vertex, and N is the number of vertexes in the mesh. In a simple mesh each vertex m_i has exactly three neighbours, whose coordinates are denoted by m_{i_j} .

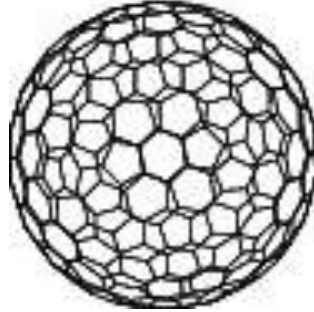


Figure 4.5 Simplex mesh with 320 vertexes. The image is from [23].

The internal energy is defined as:

$$E_{\text{int}}(m_i | m_{i_1}, m_{i_2}, m_{i_3}) = \frac{\|m_i - \frac{1}{3} \sum_{j=1}^3 m_{i_j}\|^2}{A(M)} \quad (2)$$

In (2), $A(M)$ is the average area of the faces of the mesh M . The normalization using the area is required for scale invariance of the internal energy.

The external energy is defined as:

$$E_{\text{ext}}(m_i) = 1 - \frac{(I_T(m_i))|\nabla I_T(m_i)|}{\max_x I_T(x)|\nabla I_T(x)|} \quad (3)$$

In (3), I_T is an intensity image, built from the analyzed binding potential image, by applying a threshold method. If T is the threshold, all voxels from the binding potential image having a higher value than T , will be replaced by that same threshold value. The voxel with smaller value will maintain the original intensity value. The threshold T is automatically determined for each image based on the expected volume of the striatum [3]. This threshold technique is required to correct the high inhomogeneous intensity values within the striatum caused by the partial volume effect, noise, and other image artifacts. The edge image $\nabla I_T(w_i)$ is a discrete approximation of the gradient, and it is computed using a 3-D Sobel operator [3].

The external energy defined by eq. 3 is application dependent. Therefore, any new segmentation problem requires finding a suitable equation to express the external energy. Equation 3 is used to extract the striatum surface from brain hemisphere images in [3].

For more information about the DSM algorithm see [2].

4.3.2. DSM-OS (Dual Surface Minimization-Outer Surface)

The DSM-OS algorithm (Dual Surface Minimization – Outer Surface) is a variant of the DSM (Dual Surface Minimization) algorithm. It is also initialized with two surfaces meshes: the inner surface, located inside the brain surface and outer surface, outside the brain surface (figure 4.4).

The main difference between these two deformable surface methods is that, in the DSM-OS algorithm, only the outer surface moves towards the region of interest. This is, in some cases, advantageous. For example, when extracting the brain surface from positron emission tomography (PET) images, it is preferable to approach the surface of interest from outside due to noise. Outside the brain volume there is no relevant radioactivity and therefore the noise level is lower in that region. Consequently, it is appropriate to use only one surface mesh when approaching the target surface from the outside. The algorithm optimization cycle terminates when it fulfills a condition, namely, when the outer surface reaches the inner surface. Afterwards, the algorithm remembers the mesh of the lowest energy found, and the mesh used to compute that energy is assigned as the desired segmented structure.

4.4. Incorporating “A Priori” Knowledge

In medical images, the general shape, location and, orientation of objects are often known and this knowledge may be incorporated into the deformable model as initial conditions, data constraints, constraints on the model shape parameters or, into the model fitting procedure, increasing the accuracy of the final result. The use of implicit or explicit anatomical knowledge to guide shape recovery is especially important for a robust automatic interpretation of medical images. For automatic interpretation, it is essential to have a model that not only describes the size, shape, location, and orientation of the target object but that also can cope with unexpected variations in these characteristics. Automatic interpretation of medical images spares clinicians from the laborious aspects of their work and still provides increased accuracy, consistency, and reproducibility of the interpretations [1].

5. Graph-Based Image Segmentation

This chapter describes principles of graph based image segmentation theory that are later on applied. Two types of graph image segmentation approaches are presented, namely, normalized cuts and weighted kernel k-means. The first is a spectral method and the latest a non-spectral graph based approach.

Graph-Based techniques are commonly categorized as region-based image segmentation techniques. To pose image segmentation as a graph partitioning problem, one has to represent how the chosen image features affect the relationship between voxels. This information is represented as weights of the connections between the nodes, i.e., by a weighted graph (figure 5.1) $G = (N, E)$ with node set N and edge set E [35]. The nodes $n \in N$ correspond, for example, to pixels (or voxels) intensity values or position, while the edges connect the nodes n_i, n_j according to the neighborhood definition in use. Every edge $\langle n_i, n_j \rangle \in E$ has a weight. Depending on the problem such weights might represent, for example, costs, lengths or capacities, etc.

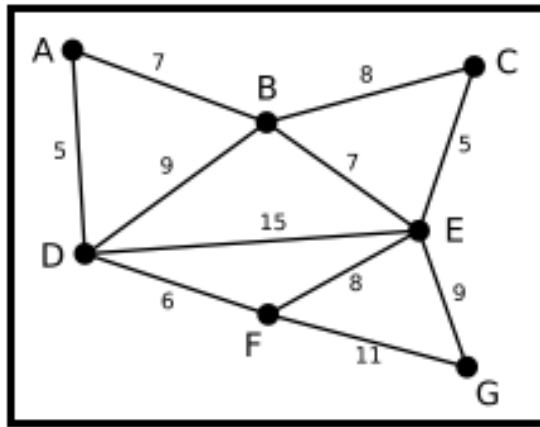


Figure 5.1 Weighted graph, where the letters are the set of nodes N , and the numbers the set of edges E .

The constructed image graphs can be directed or undirected. A directed graph is one whose edges are ordered pairs of nodes. In contrast, undirected graphs connect the nodes but the directionality of the connections is not defined. A directed graph may be thought of as a neighborhood of one-way streets. A regular two-way street may be thought of as an undirected graph [36].

Minimum spanning trees [37], shortest paths [38] and graph-cuts [39] are some of the used graph algorithms to solve image segmentation problems. Among these methods, graph-cuts is the one

that distinguishes itself, due to its computational efficiency and flexibility as a global optimization tool [35].

5.1. Spectral methods for clustering

Spectral clustering methods are region-based image segmentation methods. This approach segments different structures into distinct fragments. One variant of these methods uses the top eigenvectors of a matrix. These eigenvectors are derived from the distance between points. Such methods have been successfully used in many applications including computer vision and VLSI design [39; 40]. However, there are still debates on which eigenvectors to use and how to derive clusters from them.

One method consists in using spectral clustering to partition a spectral graph, such that the second eigenvector of the graph's Laplacian (matrix representing the image graph, computed by subtracting the degree matrix (D) and the image affinity matrix (A)) is used to define a semi-optimal cut [42]. This method assumes that cuts based on the second eigenvector will attain a good approximation to the optimal graph cut. This analysis can be extended to clustering by building a weighted graph in which nodes correspond to data-points and edges are related to the relative features of the points. Cluster differentiation is done by partitioning the graph, which can be done by a variety of algorithms such as ratio cut [43], Kernighan-Lin Objective [44], and normalized cut [45]. Usually, spectral graph partitioning objectives are used to partition the graph in two. Therefore, when one wants to partition the graph in more than two parts, the partition methods have to be applied recursively to find the k desired clusters. It has experimentally been observed [39; 40] that using more eigenvectors and the direct computation of a k way partitioning improves the results.

5.2. Normalized Cuts

As mentioned above, the normalized cuts algorithm can be used as a criterion to segment a graph. A graph can be partitioned into two disjoint sets by removing the edges connecting the two parts. One can attain the degree of dissimilarity between these two disjoint areas by computing the total weight of the removed edges [42]. This measurement is usually known in graph theory as “cut” and it can be mathematically represented as follows:

$$\text{cut}(B, C) = \sum_{u \in B, v \in C} A(u, v) \quad (4)$$

In (4), B and C are two distinct areas, while $A(u, v)$ is a function of the similarity between nodes u and v, A is also called affinity matrix.

To attain the global optimum graph bipartition one has to minimize (4). The number of partitions for a graph with N nodes is 2^N . This optimization is a well-studied problem and efficient algorithms have been developed to solve this problem. However, when there are small sets of isolated nodes in the graph, the minimum cut criteria favors the cut in those areas, even when it is not the optimal cut.

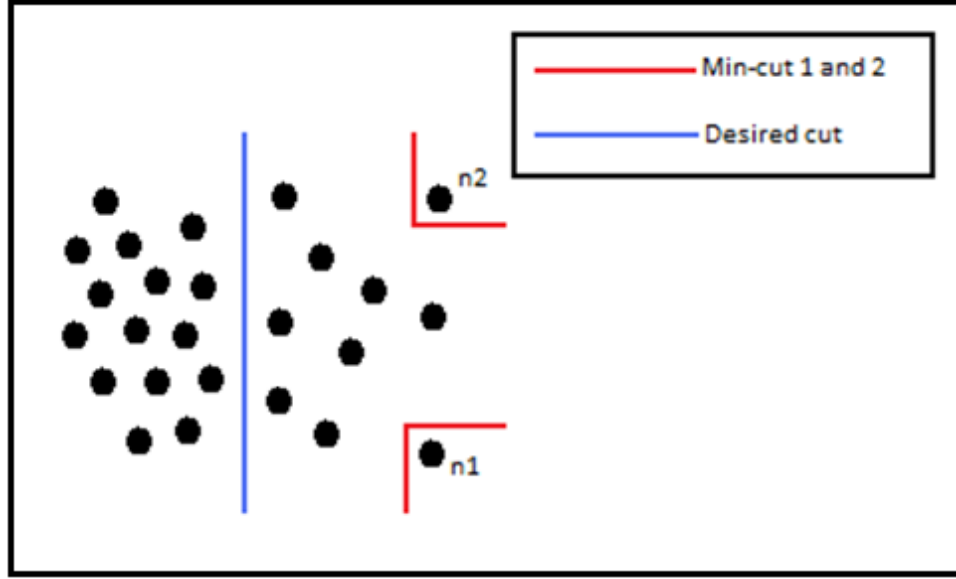


Figure 5.2 Case where the minimum cut criterion fails. The red lines represent the true cuts and the blue line is the ideal cut. The image is from [45].

In figure 5.2 is shown one case where the minimum cut criterion fails, which thereby justifies the need of more sophisticated methods to partition the graph. In this figure it is assumed that the edge weights decrease with the increase of the distance between two nodes. Therefore, when one applies the minimum cut criterion to the image, the nodes n_1 and n_2 will have the minimum cut because they are the ones furthest away from their neighbors. In fact, all the nodes on the right side of the desired cut will have smaller cut value when compared to the nodes to the left.

However, if one uses a method that, instead of relying only on the total edge weight value of the connecting partitions, it also accounts for all connections to all nodes in the graph, the problem described above can be avoided.

This new disassociation measure between two groups, named “normalized cut” (Ncut), was first introduced in [45] and is defined as:

$$\text{Ncut}(B, C) = \frac{\text{cut}(B, C)}{\text{assoc}(B, V)} + \frac{\text{cut}(C, B)}{\text{assoc}(C, V)}, \quad \text{assoc}(B, V) = \sum_{u \in B, t \in V} A(u, t) \quad (5)$$

In (5), $\text{assoc}(B, V)$ is the sum of the weights of the connections from nodes belonging to B to all nodes in the graph, and $\text{assoc}(C, V)$ is, similarly, the sum of the weights of the connections from nodes belonging to C to all nodes in the graph [45].

Using this disassociation definition between the groups eliminates the problem that arises when the minimum cut criterion tried to partition small isolated points. Therefore, using the normalized cut criterion to partition small isolated points will, in general, have a large cut value due to the denominators in (5). For example, nodes n_1 and n_2 from figure 5.2 will no longer have the smaller cut value, as they did when submitted to the minimum cut criterion. Instead, the N_{cut} value will increase because the association measure will decrease.

It is shown in [45] that the minimization of the N_{cut} criterion (5) can be converted into a generalized eigenvalue system:

$$(D - A)y = \lambda Dy \quad (6)$$

In (6), A represents an affinity matrix and D is an $N \times N$ diagonal matrix. The diagonal matrix D is attained by summing each value from each line of the image affinity matrix (A).

$$D(i, i) = \sum_j A(i, j) \quad (7)$$

The generalized eigenvalue system (6) can be transformed into a standard eigenvalue problem:

$$D^{-\frac{1}{2}}(D - A)D^{-\frac{1}{2}}x = \lambda x \quad (8)$$

Solving a standard eigenvalue problem is a difficult task since it requires high computational power. These computations take $O(n^3)$ operations to compute all the eigenvectors, where n is the number of nodes in the graph. Therefore when one applies this method to a large image, the number of operations is usually too large, due to the high number of voxels. See appendix 1 for a detailed explanation about eigenvectors and eigenvalues.

We can then partition the graph using the following algorithm [45] which goes through the following the steps:

1. Given an image or image sequence, create a weighted graph with the image information.
2. Solve $(D - A)x = \lambda Dx$ for eigenvectors with the smallest eigenvalues.
3. Use the eigenvector with the second smallest eigenvalue to bipartition the graph.
4. Decide if the current partition should be subdivided and recursively repartition the segmented parts if necessary.

In the step one, an affinity matrix is built from chosen image features. Some examples of possible features are spatial connectivity, voxel intensity values, Euclidean distance, etc.

Step two solves (8) to determine the eigenvector of the second smallest eigenvalue. The eigenvalues are the solutions of the normalized cuts mapped from a discrete domain to a continuous domain. Each component in an eigenvector corresponds to a node in the graph.

Afterwards, the graph is bipartitioned using the computed eigenvector by thresholding the eigenvector components. The search for the best threshold is done using the Ncut criterion. Only the optimum partition of the eigenvector of the second smallest eigenvalue is searched, because trying to find the optimum partition for all the eigenvectors would not be computationally feasible since there is no algorithm that can solve this problem in “polynomial time”.

Finally, in step four, the algorithm decides if the current partition should be subdivided, considering that the graph is being bipartitioned into more than two clusters. If this assumption holds, the algorithm returns to the initial step and repartitions the sub-graphs found in the first iteration.

There is another way of attaining a k-way partition of the graph that uses more than one eigenvector. If we want to partition the graph in, for example, four partitions, one uses three eigenvectors and each component of each eigenvector will correspond to an image feature. Afterwards, these features are fed to a clustering algorithm e.g., k-means, that will cluster the image based on the given features.

5.3. Non – spectral methods

Spectral clustering, after mapping the original space into an eigenspace, uses the information from the eigenvalues and eigenvectors for graph partitioning. These methods are called spectral, because they make use of the data affinity matrix spectrum to cluster the points. Their efficiency arises from not making any assumptions about the final clusters.

On the other hand, non-spectral methods do not rely on the computation of the eigenvalues and eigenvectors of the image affinity matrix.

5.3.1. Weighted Kernel k-means

The weighted kernel k-means algorithm [46] is a generalization of the standard k-means algorithm [47], and maps points to a higher-dimensional space. Therefore weighted kernel k-means has a major advantage over standard k-means, since it can discover clusters that are not linearly separable in the input space.

Weighted kernel k-means was first presented in [48]. It was shown that the weighted kernel k-means can also be used to monotonically improve a specific graph partitioning objective. However, in this case the weights and the kernel have to be chosen properly. This property allows the weighted kernel k-means to be mathematically equivalent to general weighted graph partitioning problems, such as, ratio association, ratio cuts, Kernighan-lin objective or normalized cuts [48]. This thesis focuses on the equivalence between weighted kernel k-means and the normalized cuts algorithm.

Therefore, when the eigenvector computation is prohibitive, for example, when one is analyzing large image matrixes, the normalized cuts algorithm can be replaced by the weighted kernel k-means objective.

The weighted kernel k-means tries to find clusters $\pi_1, \pi_2, \dots, \pi_n$ that minimize the objective function:

$$\mathcal{D}(\{\pi_c\}_{c=1}^k) = \sum_{c=1}^k \sum_{a_i \in \pi_c} W_i \|\phi(a_i) - m_c\|^2 \quad (9)$$

Where

$$m_c = \frac{\sum_{a_i \in \pi_c} W_i \phi(a_i)}{\sum_{a_i \in \pi_c} W_i}, \quad (10)$$

In (9), W is a diagonal matrix attained by summing each value from each line of the image affinity matrix A , its values are always non-negative. Note that the points a_i correspond to the i voxels belonging to the c -th cluster π_c , and m_c represents the “best” cluster representative:

$$m_c = \operatorname{argmin}_z \sum_{a_i \in \pi_c} W_i \|\phi(a_i) - z\|^2 \quad (11)$$

The kernel function ϕ allows the equivalence between the weighed kernel k-means and a general weighted graph partitioning objective. The kernels usually used are polynomial, gaussian, or sigmoid.

The distances are computed using inner products, since $\|\phi(a_i) - m_c\|^2$ equals

$$\phi(a_i)\phi(a_i) - 2 \frac{\sum_{a_j \in \pi_c} W_j \phi(a_i)\phi(a_j)}{\sum_{a_j \in \pi_c} W_j} + \frac{\sum_{a_j, a_l \in \pi_c} W_j W_l \phi(a_i)\phi(a_l)}{(\sum_{a_j \in \pi_c} W_j)^2} \quad (12)$$

Using the kernel matrix K , where $K_{ij} = \phi(a_i)\phi(a_j)$, (12) can be written as:

$$K_{ii} - 2 \frac{\sum_{a_j \in \pi_c} W_j K_{ij}}{\sum_{a_j \in \pi_c} A_j} + \frac{\sum_{a_j, a_l \in \pi_c} W_j W_l K_{jl}}{(\sum_{a_j \in \pi_c} A_j)^2} \quad (13)$$

To attain segmentation equivalence between weighted kernel k-means and normalized cuts algorithm, one has to assume $W = D$ and $K = D^{-1}AD^{-1}$ where W and D are diagonal matrixes attained by summing each value from each line of the image affinity matrix (A), and K is the kernel matrix. Afterwards, the weighted kernel k-means [48] can be computed following the steps below:

function $[\{\pi_c\}_{c=1}^k] = \text{Kernel_Kmeans}(K, k, W, t_{\max}, \{\pi_c^{(0)}\}_{c=1}^k)$

Input \rightarrow K : Kernel matrix ; k : number of clusters ; W : weights for each point;

t_{\max} : maximum number of iterations ; $\{\pi_c^{(0)}\}_{c=1}^k$: initial cluster.

Output \rightarrow $\{\pi_c\}_{c=1}^k$: final clusters.

1. If no initial clustering is given, initialize the k clusters (i. e., randomly).

Set $t = 0$.

2. For each point a_i and every cluster c , compute:

$$d(a_i, m_c) = K_{ii} - 2 \frac{\sum_{a_j \in \pi_c^{(t)}} W_j K_{ij}}{\sum_{a_j \in \pi_c^{(t)}} W_j} + \frac{\sum_{a_j, a_l \in \pi_c^{(t)}} W_j W_l K_{jl}}{\left(\sum_{a_j \in \pi_c^{(t)}} W_j\right)^2}$$

3. Find $c^*(a_i) = \arg\min_c d(a_i, m_c)$,

resolving ties arbitrarily. Compute the updated clusters as:

$$\pi_c^{(t+1)} = \{a : c^*(a_i) = c\}.$$

4. If not converged or $t_{\max} > t$, set $t = t + 1$ and go to step 2; Otherwise, stop and output the final

$$\text{clusters } \{\pi_c\}_{c=1}^k = \{\pi_c^{(t+1)}\}_{c=1}^k$$

6. Segmentation of striatal structures

In this chapter are described in detail the methods used here to extract the caudate, putamen and ventral striatum from high resolution PET images.

6.1. Pre-processing steps

In the pre-processing steps our purpose is to extract the striatum surface from binding potential HRRT PET (High Resolution Research Tomograph Positron Emission Tomography) images. In the figure below we can see from top to down three different HRRT PET images in the transversal, coronal and sagittal views (left-right). The technique used to achieve the segmentation of the striatum was first implemented in [23] and [3].

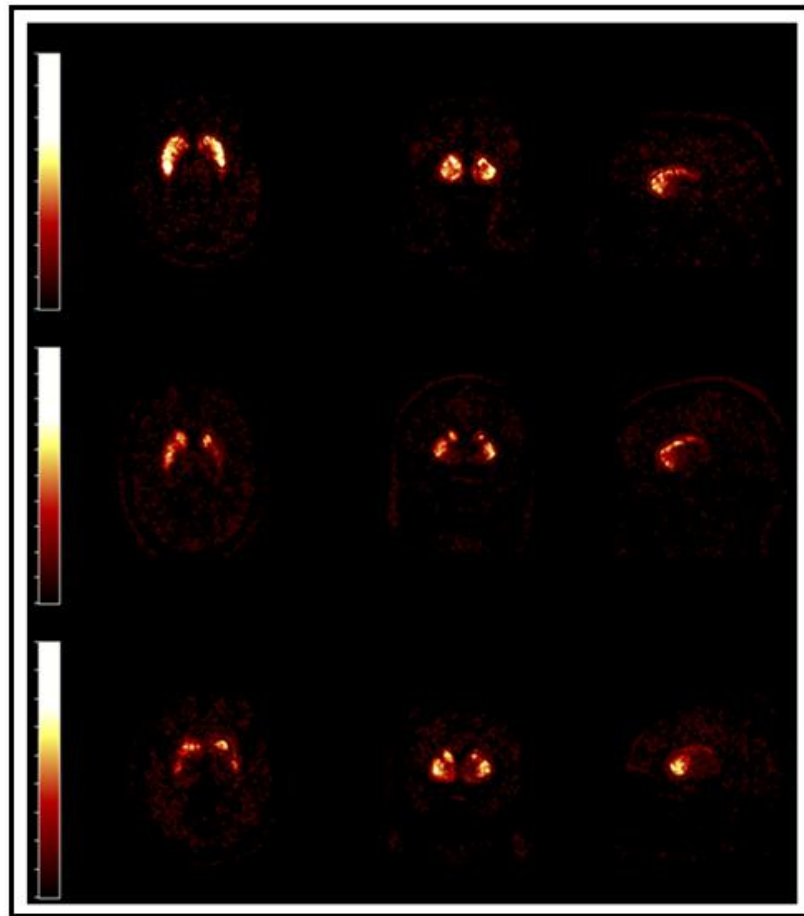


Figure 6.1 Transversal, Coronal and Sagittal views (left-right) of the segmented right striatum of three HRRT PET images (top-down), these images were image provided by Turku PET Center visualized using [49].

The pre-processing procedure can be divided in three steps: Delineation of the brain structure, segmentation of the brain structure in right and left hemisphere and extraction of the striatum from each hemisphere image. More detailed information on this method can be found in [23], [3].

The delineation of the brain structure was achieved using a deformable model which is based on global minimization of its energy. In other words, the algorithm searches for the surface that minimizes the energy function $E(;\cdot)$ of the image. Given an image I , the energy $E(S;I)$ of the surface S depends on the image data and the properties of the surface. The total energy of the surface is a weighted sum of two energies, the internal energy that constrains the shape of the surface penalizing surfaces that are not smooth, and the external energy that combines the surface with the salient image features. The energy function (1) was minimized with the global DSM-OS (Dual Surface Minimization – Outer Surface), this algorithm is a variant of DSM (Dual Surface Minimization). DSM-OS was used because sometimes it is advantageous to approach the surface of interest from a specified direction, in this case it is from outside because outside there is considerable less noise than from inside the brain structure. In this algorithm only one of the two surfaces is allowed to move and the other one is used only for computing the stopping condition. Then, the algorithm just remembers the surface of lowest energy that has already been found and compares with surfaces found in subsequent iterations of the algorithm to it, as explained in chapter four.

After having extracted the brain surface we have to divide that surface in to hemisphere images (figure 6.2). This is done using an automatic method called “mid-sagittal plane extraction”. This method is very similar to an existing edge-based method originally designed for anatomical brain images [50]. The difference from this algorithm and the existing one is that in this one it is applied to the extracted brain structure instead of the edge-images due to the low contrast to noise ratio in PET images. The mid-sagittal plane was defined as the one about which the reflective symmetry of the brain is maximized.

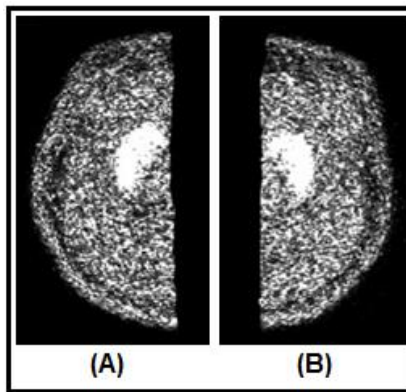


Figure 6.2 (A) left brain hemisphere, (B) right brain hemisphere.

The last computation of this pre-processing step is the segmentation of the striatum from the two hemispheres images, this is achieved using DSM-OS algorithm like in the brain surface extraction. The main difference between the DSM-OS algorithm used in the brain delineation and in the striatum extraction is in the initial surfaces. In the striatum extraction the initial surfaces are generated automatically by first locating a voxel within (left/right) striatum based on the centre of mass of the highest BP values in a (left/right) hemisphere image and then centering a large enough sphere surface to that voxel, while in the brain surface extraction one places one initial surface inside the brain area and the other outside, this does not require any special algorithm because of the known brain geometry.

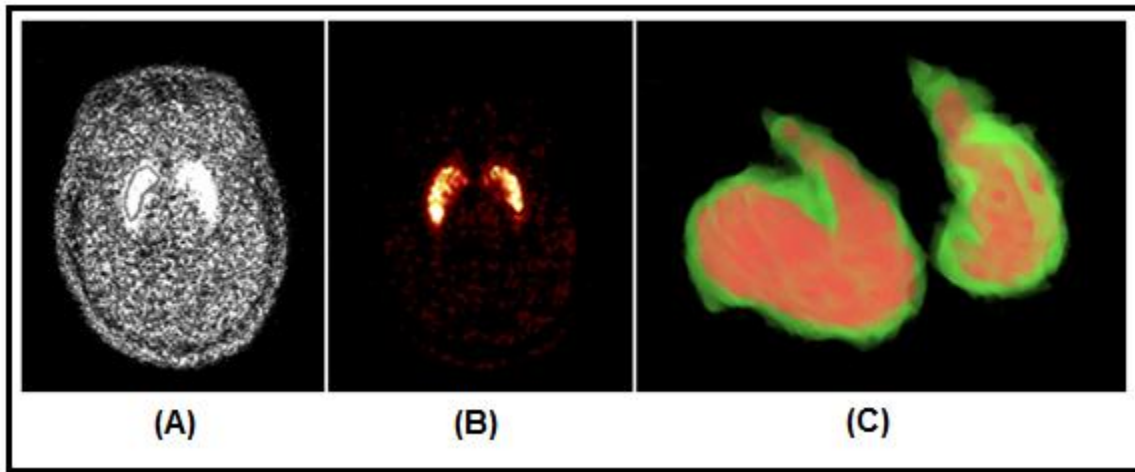


Figure 6.3 (A) segmented left striatum, (B) segmented right striatum, (C) 3D render from segmented striatum's. The image on the left and center were taken from slice number 77 of the original image; (A) and (B) were reproduced using [51] and (C) using [52].

6.2. Segmentation of the Striatum into four structures

Our method for the segmentation of the striatum into four distinct structures proceeds in a stepwise fashion, taking into account the physical and spatial anatomy of the structures of interest, the large variations of voxel intensity values within and between the structures, and the existence of a thin plate of white matter, with relatively smaller voxel intensity values, separating the Putamen and the Caudate.

Once the striatum is segmented in the pre processing steps we have the information of which voxels belongs solely to the striatum, something that allows us to apply the rest of the method only to a small portion of the real image size, this enhances the speed, robustness and effectiveness of the procedure.

We build the affinity matrix also known as weight matrix using only the voxels belonging to the striatum, we define the affinity matrix as (A) and this matrix describes the inter-voxel affinity (intensity values and voxel distances) that is represented as weights between voxels, this information allows us to know the features needed to the segmentation of the image into different structures.

For an image I with dimensions i, j, k, respectively on the x, y, z axis, where i, j, k $\in \mathbb{N}$, its affinity matrix A has a dimension of i * j * k rows and i * j * k columns such that A(r, :), r $\in \mathbb{N}$ is a vector of similarity between the measured features from voxel I(r, r) and all the other voxels in the image I, e.g. figure 1.3. Analyzing the values of the affinity matrix we conclude that if x $\in A$ then x $\in [0,1]$, and the closest the value of x is to 1, more similar x becomes to the voxel corresponding to the line being analyzed.

Affinity between two voxels can be qualitatively classified by comparing the difference between the two voxel intensity values with a given maximum threshold value and their spatial location. If the difference exceeds a chosen threshold, the two voxels are assumed to belong to different classes or, more precisely, different structures. The pairwise differences between the voxels are mapped into the affinity matrix, and the clustering algorithm analyzes these affinities, clustering voxels into structures according to their affinity degree.

Once the affinity matrix is constructed the affinity degrees are used to assign each voxel to its correct structure.

6.2.1. Affinity matrix construction methods

The voxel affinity matrix was built to deliver accurate information about voxel relation to the segmentation method. Two different approaches were used. The first one was first implemented by Shi and Malik in [45] and the second is a thresholding method developed in this study.

Shi and Malik's used two features to build the affinity matrix (A), intensity values similarity and Euclidean distance between voxels. The Euclidean distance (Dist) between points x = (x₁, x₂, ..., x_n) and y = (y₁, y₂, ..., y_n), in an n dimensional space is given by:

$$\text{Dist}(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (14)$$

The weight/affinity between node i and j is A_{i,j}, and can be computed as:

$$A(i, j) = e^{-|I(i) - I(j)|} * \begin{cases} e^{-\text{Dist}(i, j)} & \text{if } \|X(i) - X(j)\|_2 < r \\ 0 & \text{otherwise,} \end{cases} \quad (15)$$

In (15) the two terms of the equation represent, respectively, the similarity between the voxel i and j intensity values and the Euclidean distance between voxel i and j spatial location. Notice that if the Euclidean distance between voxel i and j is above a given threshold, the affinity degree is set to zero.

On the other hand, the method developed in this thesis gives better results than Shi and Malik method when applying the affinity matrixes from both methods to the same segmentation problem. The method is depicted below:

$$A(i, j) = \begin{cases} e^{\text{norm} * |I(i) - I(j)|} & \text{if } I(i) \text{ is connected to } I(j) \\ 0 & \text{otherwise,} \end{cases} \quad (16)$$

This method assumes that voxels $I(i)$ and $I(j)$ are connected if the distance between them is below the threshold. The variable norm is a normalization parameter of the intensity of each voxel, mapping all values between -1 and 0. Once the affinity matrix is obtained, accounting for the similarity between intensity values (16), another threshold is applied (17). This threshold removes connections between any two voxels with a degree of intensity similarity below the threshold, enhancing structure boundaries and allowing a more accurate clustering.

$$A(i, j) = \begin{cases} A(i, j) & A(i, j) < \text{threshold} \\ 0 & \text{otherwise,} \end{cases} \quad (17)$$

In (17) $A(i, j)$ corresponds to the affinity value between voxel i and j .

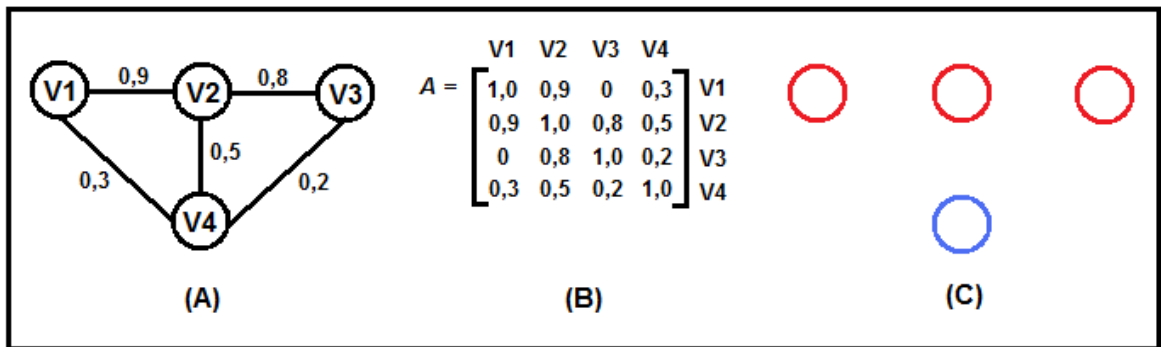


Figure 6.4 (A) Graph representing an image with four voxels (V1, V2, V3, V4) and five weighted edges, (B) affinity matrix build using image (A) information, (C) clustering result of image (A) using a 2-way segmentation algorithm.

In figure 6.4 there is an example of a four node weighted graph (A), each node represent one image pixel, in the (B), there is an example how the graph affinity matrix will look like, in (C), one can find an example of the clustering when the affinity matrix (B) is feed to an clustering algorithm.

6.2.2. Clustering methods

Two clustering techniques are here implemented, the first being a “multiway normalized cuts” algorithm [45] and the second being the “weighted kernel k-means” algorithm [48].

The multiway normalized cuts that makes use of eigenvectors, i.e., is a spectral method. The first step of this algorithm is constructing an affinity matrix by assigning a node to each voxel in the image, and an edge weight for all pairwise neighboring voxels. The assigned weight value is a function of the level of matching between the features of the two voxels. This formulation allows treating the image segmentation problem as a graph partition of the vertex set. The method seeks to partition the graph into subgraphs, internally highly connected, but weakly connected between one another, using the computed first k eigenvectors of the graph Laplacians to define the feature space to which the K-means clustering algorithm is applied.

The K-means algorithm is an unsupervised learning clustering algorithm. The procedure classifies a given data set making use of a defined number of fixed clusters. The algorithm starts by defining k -centroids, one for each cluster, and then takes each point belonging to the data set and associates it to the nearest centroid. The first step is completed when all points are assigned to one of centroids. Next, k new centroids are chosen, and correspond to the mean value of the data points assign to each previous k cluster. A new clustering assignment between the same data set points and the nearest new centroids is then done. These steps are repeated until there are no more changes in the centroids location. The algorithm aims to minimize the total intra-cluster variance, or, the squared Euclidean distance objective function:

$$J = \sum_{c=1}^k \sum_{a_i \in \pi_c} \|a_i - m_c\|_2^2 \quad (18)$$

In (18) there are k clusters π_c , $c = 1, 2, \dots, k$ and m_c is the centroid or mean point of the points a_i belonging to the cluster π_c .

function [Clustered] = Multiway Normalized Cuts(Image, k)
 Input → Image: image that we want to be segmented, k: number of clusters.
 Output → Clustered: Image with the final clusters.

1. Build the affinity matrix;
2. Compute the k first eigenvectors of $(D - A)y = Dy$;
3. Use K_{Means} algorithm to cluster the k eigenvectors into clusters C_1, \dots, C_k ;

Considering $W = D$ and $K = D^{-1}AD^{-1}$ where W and D are diagonal matrixes attained by summing each value from each line of the image affinity matrix (A). The new affinity matrix K is positive definite (positive definiteness guarantees convergence).

The weighted form of the kernel k-means is mathematically equivalent to a general weighted graph clustering. Such equivalence implies that the weighted kernel k-means algorithm can be used to locally optimize a number of graph clustering objectives such as normalized cuts. This is done generalizing the k-means objective function to use both weights and kernels [48] from which the altered distance function (equation 13) is computed.

Therefore in cases where eigenvector computation is prohibitive, for example, if many eigenvectors of a large matrix are required, the weighted kernel k-means algorithm is a more suitable choice than spectral methods. The algorithm described below is a representation of a clustering method using the weighted kernel k-means as the objective that partition the image into connected subsets.

function [Clustered] = Algorithm1 (Image, k, init_{clusters})
 Input → Image: image that we want to be segmented, k: number of clusters,
 init_{clusters} : initial clusters .
 Output → Clustered: Image with the final clusters.

1. Build the affinity matrix;
2. Use Weighted Kernel K_{Means} to cluster the image affinity matrix
 into clusters C_1, \dots, C_k ;

6.3. Implementation

In this section are described the methods implemented in this thesis. After tuning the initial parameters the method implemented in [3] was applied to extract both right and left striatum from the high resolution PET images. In [3], the authors used a 2-way normalized cut algorithm to segment the two structures that they proposed to extract (caudate and putamen). However, this thesis aim is to extract four structures (caudate, putamen, ventral striatum and white matter), so a k-way normalized cut algorithm was developed. After realizing that the multiway normalized cuts algorithm computation failed when applied to the high resolution PET images, a new approach was developed. This approach aimed to replace the normalized cuts, giving a similar segmentation results but applicable to high resolution images. The developed algorithm was based on the weighted kernel k-means [48].

Both segmentation algorithms have to be fed with a voxel feature matrix (affinity matrix). This affinity matrix is built using a set of chosen features that are extracted from the image voxels. The construction method of the affinity matrix is one of the most important steps on the segmentation of the striatal structures, because it is in the matrix that is stored the information that will be used in the segmentation process. In this thesis were implemented two construction methods of the affinity matrix, the one proposed by Shi and Malik in [45], and the one build by us. Our method exhibited, when applied to the created phantoms, more accuracy and robustness to noise than the Shi and Malik method. All the created methods were implemented using Matlab [51].

7. Experiments

7.1. Materials

The evaluation of the validity of the automatic segmentation method was achieved using dynamic two dimensional phantom images and four [^{11}C] Raclopride BP (Binding potential) HRRT PET images.

The phantoms used to test the methods are shown in figure 7.1. Figure 7.1 (B) is a 2D phantom with dimensions 32x32. This phantom is a simple representation of the structures that this thesis aims to extract. In this image the yellow region corresponds to caudate, the dark blue to putamen, the light blue to white matter and the dark red to the ventral striatum.

The figure 7.1 (A) is a “colormap” taken from the coronal view, slice 112, of one of the [^{11}C] Raclopride HRRT PET scans. This phantom allows measuring the reliability and reproducibility of the segmentation method. It also allows understanding the correlation between different voxel intensity values within the same structure (cluster), as well as the segmentation smoothness in those structures boundaries.

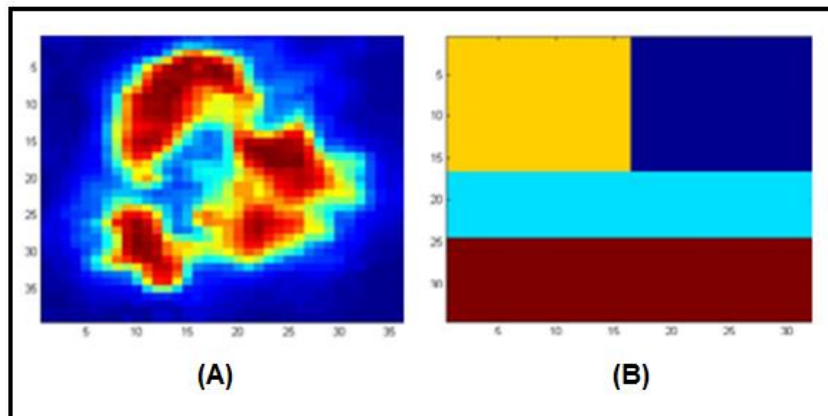


Figure 7.1 (A) 39x36 Phantom, (B) 32x32 phantom.

The [^{11}C] Raclopride HRRT PET scans were provided by Turku PET center and were from four healthy male volunteers. The subjects were free from any somatic or psychiatric illness, drug abuse and alcoholism.

The imaging was attained by a brain-dedicated, high-resolution PET scanner (ECAT HRRT, Siemens Medical Solutions, Knoxville, TN) [53]. In the acquired measurements, a molded thermoplastic mask was used as a head fixation to reduce the noise due to head movements.

The PET-raclopride scans were acquired in list mode and histogrammed by axial compression of span 9. The raclopride uptake was measured for fifty one minutes after injection using thirteen frames. The frame sequence consisted of three frames of sixty seconds, four frames of one hundred and eighty seconds, and six frames of three hundred and sixty seconds. The image size of each scan was $200 \times 200 \times 150$ slices and they have an isotropic voxel dimension of $1.22 \times 1.22 \times 1.22 \text{ mm}^3$. The raw data was reconstructed with a speed-optimized version of OP-OSEM-3D (Ordinary Poisson-OSEM in full 3D) [54], with sixteen subsets and eight full iterations [55]. The parametric binding potential HRRT PET images were created from the dynamic images using the simplified reference tissue model [25].

7.2. Results

In this section we present the results attained using the phantoms and real data. The phantoms were used to test and validate the proposed methods. Moreover this test allows tuning the parameters of the model. Using the tuned model, the methods were applied to the real data.

7.2.1. Phantoms

The dynamic 2D phantoms are now used to compare the accuracy of the two methods of attaining the affinity matrix by submitting the results of each method to the same clustering algorithm, namely, the multiway normalized cuts.

In figure 7.2 each line represents a level of noise added to the original image. Respectively, from top to bottom the noise level added is 0%, 10%, 30%, 60%. On the first row is represented the original data set of test images (figure 7.2 (A), (B), (C), (D)). On the second, the segmentation result using Shi and Malik's affinity matrix construction method (figure 7.2 (E), (F), (G), (H)), and on the third row are the results of our method, (figure 7.2 (I), (J), (K), (L)).

Both methods gave similar results in noiseless conditions when clustering the data into four clusters (figure 7.2 (A),(E),(I)) but, when random noise was added to the phantoms, the Shi and Malik's method was less robust to noise when compared to our method.

The results shown were obtained assuming that two voxels are neighbors if the difference between their spatial locations does not exceed 10 voxels. The threshold intensity value used was 0.4, i.e. if two voxels have a similarity degree below 40%, their intensity degree is set to zero, thus forcing them to belong to different clusters.

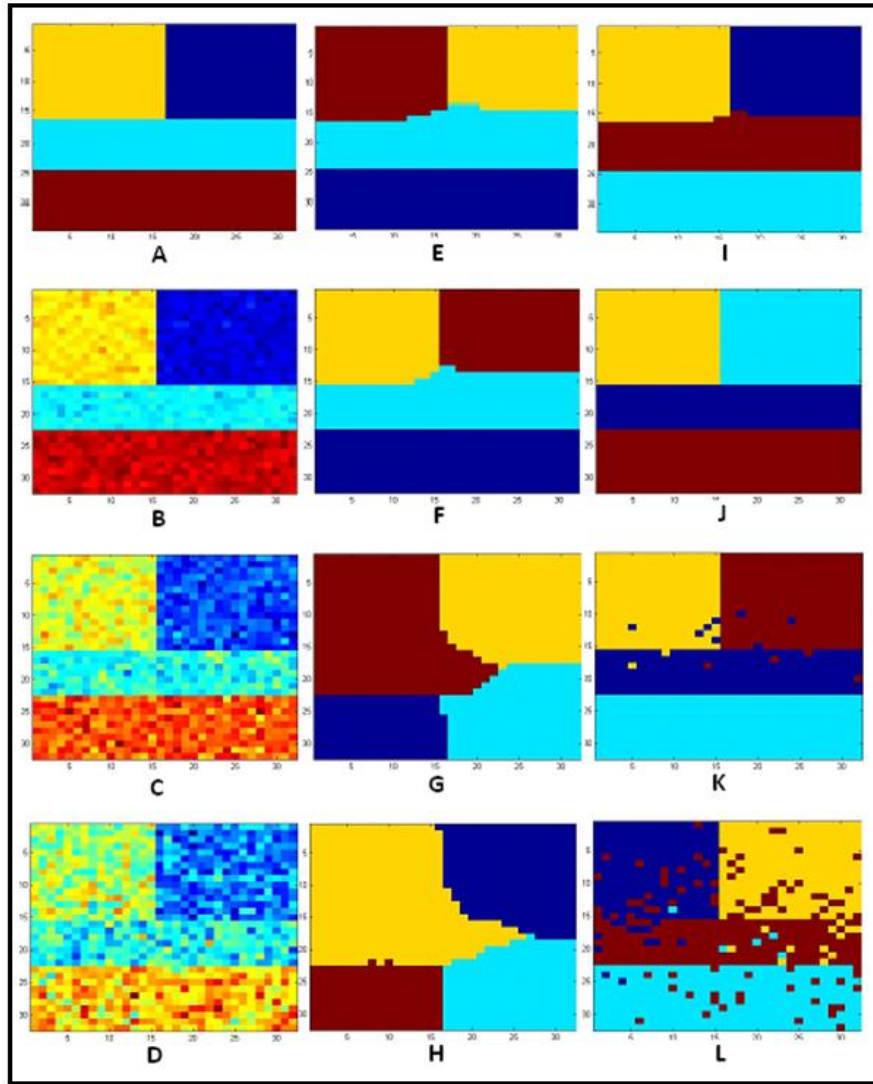


Figure 7.2 (A,B,C,D) Original images with top-bottom 0%, 10%, 30%, 60% of added random noise, (E,F,G,H) results of the segmentation of the original images using multiway normalized cuts with spatial connectivity of 10 voxels and Shi and Malik method of building the affinity matrix, (I,J,K,L) results of the segmentation of the original images using multiway normalized cuts with spatial connection of 10 voxels and a intensity threshold of 0.4 on the affinity matrix. Note that we are analyzing the structure topology, therefore the colors of the final clusters are not part of the final result. The clusters colors differ because our method does not specify any specific color for the final clusters, so Matlab chooses randomly the color the final clusters.

The phantom from figure 7.3 was used to measure the reliability and reproducibility of the construction methods of the affinity matrix. In the illustration on the left (original phantom) one can distinguish 4 main regions: the background (dark blue), the different striatal structures (red and orange) and the outline of the structures (light blue). The construction methods of the affinity

matrix proved again to be similar, i.e., had similar segmentation results using the multiway normalized cuts algorithm.

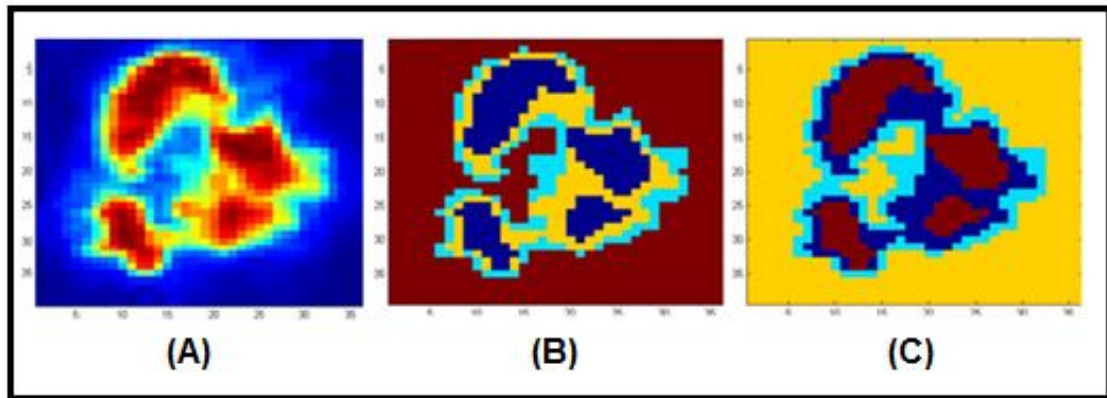


Figure 7.3 (A) Original image, (B) segmented image using multiway normalized cuts with spatial connectivity of 15, (C) segmented image using our affinity matrix construction method using spatial connectivity of 15 and a similarity threshold of 0.4.

In figure 7.4 are shown the results of testing different values of spatial connectivity keeping the similarity threshold constant and equal to 0.9. The clustering was done by the multiway normalized cuts algorithm. In this case, the value of spatial connectivity used is 15 (figure 7.4 (E)), found to be more suitable to analyze the anatomy of the desired ROIs (Regions Of Interest).

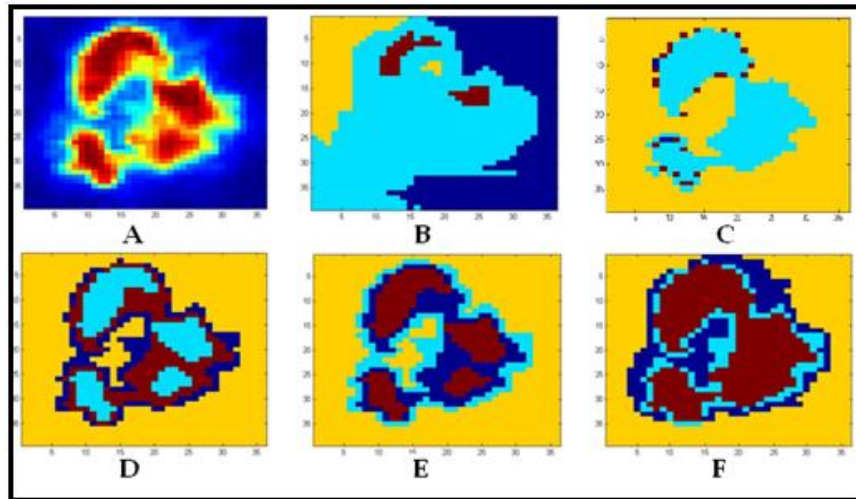


Figure 7.4 Original image (A), all the other images are the segmentation result of the multiway normalized cuts algorithm using our method for building the affinity matrix with fixed similarity threshold of 0.9 and variable spatial connectivity parameter, (B) spatial connectivity parameter (sc) of 1, (C) sc = 5, (D) sc = 10, (E) sc = 15, (F) sc = 20.

Keeping the spatial connectivity constant (at 15 voxels), the best segmentation results are obtained using a high threshold of 90%, that allows keeping only the most similar voxel values of the affinity matrix (figure 7.5 (D)).

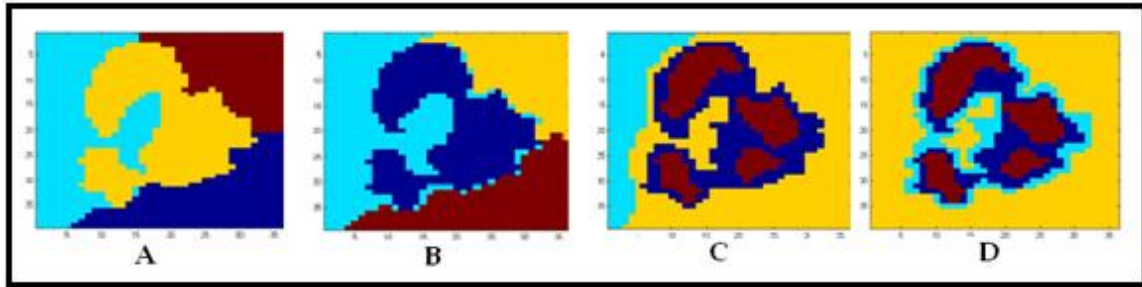
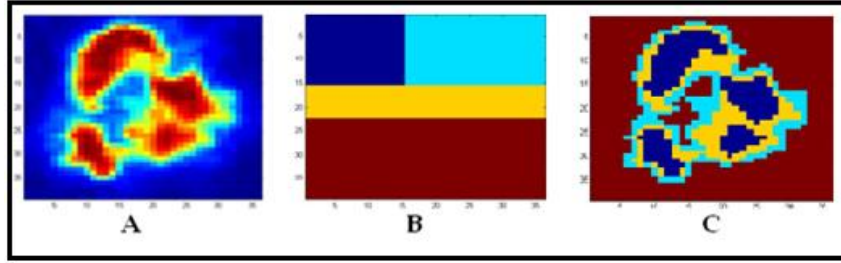


Figure 7.5 Segmentation result of the multiway normalized cuts algorithm using our method for building the affinity matrix with fixed spatial connectivity parameter of 15 voxels and variable similarity threshold, (A) using no threshold, (B) threshold = 0.3, (C) threshold = 0.6, (D) threshold = 0.9.

Although the spectral method for segmentation gives good results, it is well known that for large data sets it becomes computational prohibitive to compute the Eigenvalues of the affinity matrix. Therefore, a weighted version of the kernel k-means algorithm was used. This algorithm does not rely on the computation of eigenvalues to achieve the final clustering. This method was shown to be equivalent to the multiway normalized cuts in [48] and the results in figure 7.4 (D) (multiway normalized cuts) and figure 7.6 (C) (weighted kernel k-means) demonstrate that both methods lead to similar segmentation results. However the weighted kernel k-means achieves the segmentation result with less computational requirements when compared to the multiway normalized cuts. The weighted kernel k-means algorithm applied to the phantom of figure 7.6 is not sensitive to its initial cluster guess. Therefore we can make a random initial guess for the clustering result or use the cluster shown in figure 7.6 (B), still having the same segmentation result as in figure 7.6 (C). On the other hand, the spatial connectivity between voxels and the similarity threshold are highly sensitive to the initial guess. The best results obtained for the phantom in (figure 7.6 (A)) are shown in (figure 7.6 (C)) where it was predefined a neighborhood of 30 voxels, i.e. all the surrounding voxels at a distance smaller than 30 voxels are assumed neighbors of the voxel in question. The similarity threshold used to compute figure 7.6 (C) was 0.5, i.e. voxels with intensity similarity equal to or higher than 50% would be maintained, while the rest would be replaced by zero.



Figures 7.6 (A) original image, (B) initial cluster, (C) final clustering using in the weighted kernel k-means algorithm with spatial connectivity parameter of 30 voxels and threshold of 0.5.

7.2.2. Real HRRT images

In figure 7.7 are shown two histograms of the intensity values within the left and right striata. The test data used to build these histograms was a HRRT PET scan. As mentioned above the extraction method considers that there are four structures inside the striatum. Analyzing the histograms presented in (figure 7.7 (B), (C)), that correspond respectively to the left and right striatum intensity values. One can see that the intensity values of the several structures overlap. This overlapping makes it impossible to distinguish them if relying only on the information extracted from the intensity values. For this reason, the segmentation process of the structures that compose the striatum is a complex process.

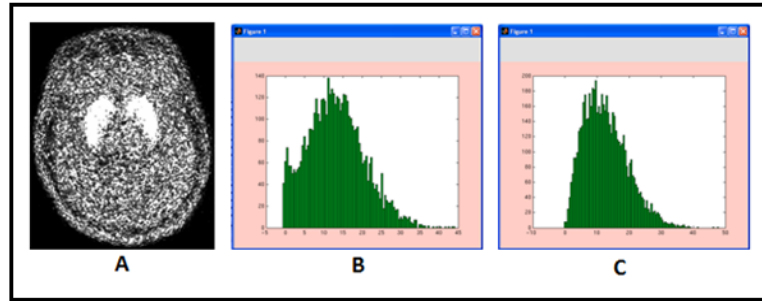


Figure 7.7 Histogram of the intensity values of both striata of the HRRT PET image; (A) original image, (B) histogram of the left striatum, (C) histogram the right striatum.

Due to this complexity, robust methods had to be developed to deal with the extraction of the striata from each HRRT PET scan. The first method that was applied was the multiway normalized cuts that due to the data size proved to be computationally too expensive, in terms of memory and time, as predicted.

Afterwards, we applied the weighted kernel k-means using our affinity matrix construction method because it is more resistant to noise and the HRRT PET images are always noisy.

Figure 7.8 shows the result of applying the weighted kernel k-means algorithm to a real HRRT PET image. The segmentation results are satisfactory, namely, the algorithm found four clusters, and from those four clusters one can easily recognize three structures: putamen, nucleus accumbens (ventral striatum) and, a merge of caudate and white matter. The green and red clusters seen in figure 7.8 (B) correspond, respectively, to the posterior and the anterior putamen (see appendix 3, yellow cluster is the posterior putamen and white cluster is the anterior putamen), the blue cluster corresponds to the caudate and, the yellow cluster corresponds to the ventral striatum. The putamen with this method was segmented into two clusters due to its highly non-uniform architecture and the existence of a large tract of white matter traversing through the structure. The white matter was the only surface of interest that was not well segmented. In all segmented images, the white matter always came embedded into other cluster.

Although in the phantom experiments the initial clustering was irrelevant, the same cannot be said when the weighted kernel k-means algorithm was applied to HRRT PET scans, because the final result depends on the initialization. The best segmentation results were obtained when the initial cluster was separated into four clusters where each cluster was a horizontal parallelepiped with the same area. When we tried different initial clusters, such as a random initialization, the result was a single cluster on the final segmentation (figure 7.8 (D)). The spatial connectivity parameter needs also to be “tuned” to achieve the best segmentation. In our case, a neighborhood of 3-4 voxels achieved the best results, while values below gave a final segmentation equal to the initial cluster and superior values resulted in a single final cluster (figure 7.8 (C), (D)). This method always segmented the image into four clusters regardless of increasing the initial number of clusters set beyond four (figure 7.8 (B)). See appendix 2 for more results on the segmentation of the four HRRT PET images into four clusters.

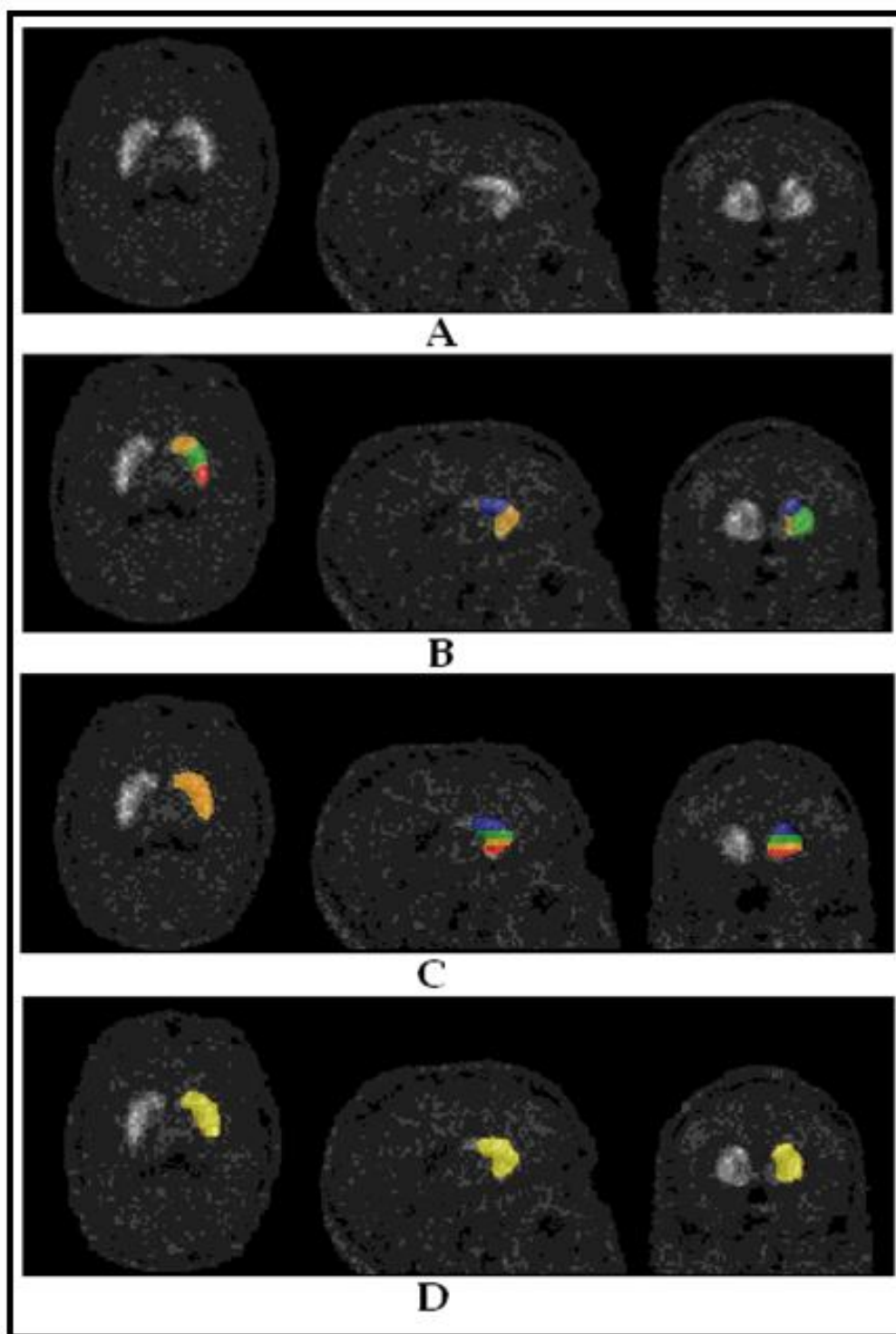


Figure 7.8 (A) Original HRRT PET BP image (B) segmented image using weighted kernel k-means with spatial connectivity of 4 and the initial number of clusters set to 4 with equal length, (C) segmentation result using spatial connectivity 1, (D) segmentation result using a random initial clustering or spatial connectivity parameter above 4. The images are Sagittal and transverse views from slice 90, and coronal view is from slice 85;

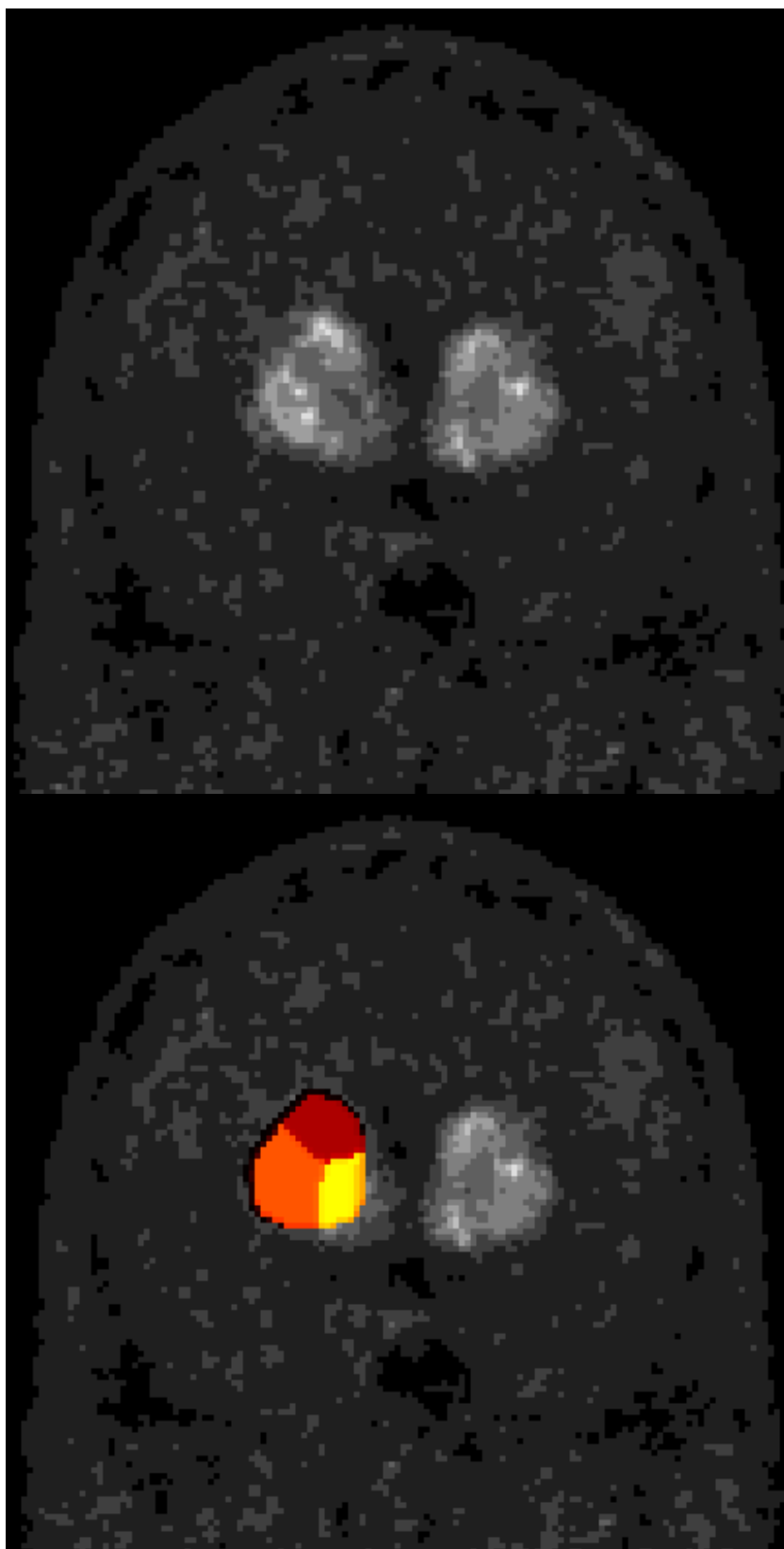


Figure 7.9 (Top) Original HRRT PET image (Bottom) segmentation result of the left striatum using weighted kernel k-means algorithm.

8. Discussion and conclusion

A new automatic method was proposed in this study. The method aims to extract striatal structures from HRRT PET binding potential images. A similar approach was done for the previous scanner model (ECAT EXACT HR+, Siemens CTI, Knoxville, TN) in [3].

The images were converted into graphs and thus, the segmentation process became a graph partition problem. It was experimentally confirmed that a multiway version of the spectral segmentation algorithm (normalized cuts) fails when applied to PET images acquired using the ECAT HRRT scanner. To solve this problem, the normalized cut algorithm was replaced by a non-spectral clustering method called “weighted kernel k-means”, giving similar segmentation results, but requiring far less computational resources.

Weighted kernel k-means does not rely on the heavy computation of the eigenvalues and eigenvectors of the image affinity matrix. It can be categorized as a region-based segmentation method that iteratively groups neighboring voxels accordingly with their similarity degree. This properties and the mathematical equivalence between the results of both algorithms makes the new algorithm more appropriate to be applied in images with considerable size, where the computation of eigenvalues and eigenvectors becomes infeasible.

The method to construct the image graph (affinity matrix) was also changed, as to improve the results previously obtained using the method proposed in [45]. Namely, after building the graph using spatial connectivity and image intensity values as features, an intensity threshold technique is applied to that graph. This improvement produces better clustering results, as well as better smoothness control over the clustering results.

The combination of these two methods, namely, the affinity matrix construction method and the weighted kernel k-means clustering algorithm, made possible the accurate extraction of the putamen, caudate and ventral striatum from four different patients scans. Moreover, the method was able to divide the putamen into two different clusters (anterior and posterior putamen) in every experiment, a result that was not initially expected but that proved to be accurate. Jussi Hirvonen, MD, PhD from Turku PET Center, University of Turku/Turku University Central Hospital verified and agreed with the accuracy of this segmentation results. For more information about this separation see [56].

The only structure that could not be extracted with this method was the white matter (internal capsule) within the striatum. The white matter appeared merged with the caudate complex in the final clustering. One way to overcome this problem could be to separate the final caudate-white matter cluster from the rest of the clusters and apply again the weighted kernel k-means algorithm

to achieve their separation. Another possibility would be to improve the accuracy of the guesses of the initial clusters. All the initial clustering guesses used in the weighted kernel k-means were simple and did not regard any anatomical information. For example, when trying to segment the striatal area of one HRRT brain image into four clusters, the initial clustering guess was an image, with the same size of the original image, divided into four areas with the same size, or a random clustering initialization. The improvement of these initial guesses is possible if one accounts for the real anatomical position of these structures that has been experimentally detected. The improvement of the initial guesses would make the clustering algorithm far more complex but, on the other hand, would improve the final clustering result and speed of the algorithm.

The striatum and its structures are important brain structures when studying the binding on dopamine D_2 receptors. Therefore, a fully automatic method for the extraction of these structures benefits brain research. This method proved to be accurate and reproducible, making it suitable to be used for multi-center preclinical and clinical trials for drug development. Another possible application for this method would be in medical procedures, where the segmentation is still done manually or semi-automatically. This would require a fast performance. The method here proposed takes only a few minutes to achieve the clustering result implemented in Matlab [51], so, its implementation using a low level programming language would further increase its speed.

These methods for automatic extraction of volumes could be further developed for other medical functional imaging technologies, other anatomical structures and radioligands (tracer). Furthermore, with the appearance of new scanners with improved resolutions, more structures of interest can be extractable from the scans, e.g. in the striatum, instead of the current four extractable structures, one could extract more substructures. The increase of the desired number of structures one wants to extract from the scans would not affect the algorithm developed in this study, since it was built to cope with increasing complexity. Moreover, automatic methods will become even more necessary, because the increase of resolution implies an increase of data size and number of slices, making the manual ROI extraction harder.

References

- [1] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: a survey. Medical Image Analysis", 1(2):91-108, 1996.
- [2] J. Tohka, "Global Optimization – Based Deformable Meshes for Surface Extraction from Medical Images", Ph.D. dissertation, Tampere University of Technology, Finland, 2003.
- [3] Tohka J., Wallius E., Hirvonen J., Hietala J., Ruotsalainen U., "Automatic Extraction of Caudate and Putamen in ^{11}C Raclopride PET Using Deformable Surface Models and Normalized Cuts", IEEE Transactions on Nuclear Science, vol. 53, Issue 1, pp. 220 – 227, 2006.
- [4] Webster's New World Medical Dictionary, 2nd Edition by MedicineNet.com, January 2003.
- [5] Nolte, J, "Human Brain: An Introduction to Its Functional Anatomy", publisher. Mosby vol. 5, book, 2001.
- [6] Packard MG, Knowlton BJ, "Learning and memory functions of the Basal Ganglia", Annu Rev Neurosci 25:563-593, 2002.
- [7] J. Crinion, R. Turner, A. Grogan, T. Hanakawa, U. Noppeney, J. T. Devlin, T. Aso, S. Urayama, H. Fukuyama, K. Stockton, K. Usui, D. W. Green, C. J. Price, "Language Control in the Bilingual Brain", Science, Vol. 312., n. 5779, pp. 1537 – 1540, 2006.
- [8] Pitman, "Functions of the basal ganglia", Ciba foundation symposium 107, London, 1984.
- [9] Wayne C. Drevets, Clara Gautier, Julie C. Price, David J. Kupfer, Paul E. Kinahan, Anthony A. Grace, Joseph L. Price, and Chester A. Mathis, "Amphetamine-induced dopamine release in human ventral striatum correlates with euphoria", Biological Psychiatry, Vol. 49, Issue 2, pp 81 – 96, 2001.
- [10] G. Berding, Th. Brcke, P. Odin, D.J. Brooks, H. Kolbe, P. Gielow, H. Harke, B.O. Knoop, R. Dengler, W.H. Knapp, "[123I]-CIT SPECT imaging of dopamine and serotonin transporters in Parkinson's disease and multiple system atrophy", Nuklearmedizin, Schattaue, Vol. 42, pp. 31 – 38, 2003.

-
- [11] Voorn P, Vanderschuren LJ, Groenewegen HJ, Robbins TW, Pennartz CM, “Putting a spin on the dorsal-ventral divide of the striatum”, *Trends Neurosci*, 27:468-474, 2004.
 - [12] Alexander GE, Crutcher MD, DeLong MR. “Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, “prefrontal” and “limbic” functions”, *Prog Brain Res*, 85:119-146, 1990.
 - [13] Heimer L. “Basal forebrain in the context of schizophrenia”, *Brain Res Rev*, 31:205-235, 2000.
 - [14] Robbins TW, Everitt BJ. Limbic-striatal memory systems and drug addiction. *Neurobiol Learn Mem*. 2002;78:625-636.
 - [15] Fudge JL, Haber SN, “Defining the caudal ventral striatum in primates: cellular and histochemical features”, *J Neurosci*, 22:1078-1082, 2002.
 - [16] Parkinson JA, Olmstead MC, Burns LH, Robbins TW, Everitt BJ, “Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine”, *J Neurosci*, 19:2401-2411, 1999.
 - [17] Maldonado-Irizarry CS, Kelley AE, “Excitotoxic lesions of the core and shell subregions of the nucleus accumbens differentially disrupt body weight regulation and motor activity”, in *rat. Brain Res Bull*, 38:551-559, 1995.
 - [18] Henk J. Groenewegen, “The ventral striatum as an interface between the limbic and motor systems”, *the international journal of neuropsychiatric medicine (CNS Spectr.)*, pp.887-892, 2007.
 - [19] Tony F. Chan, Hongwei Li, Marius Lysaker and Xue-Cheng Tai, “Level set method for positron emission tomography,” *International Journal of Biomedical Imaging*, pp. 1 – 21, 2007.
 - [20] University of California, UCLA School of Medicine, Institute for Biological Imaging, “Let’s Play PET”, 1993.
 - [21] Gerrit Westera and P. August Schubiger, “Functional Imaging of Physiological Processes by Positron Emission Tomography,” *News Physiol Sci* 18: 175-178, 2003.

-
- [22] A. P. Happonen and S. Alenius, "A Comparison of Sinogram and Stackgram Domain Filtering Methods Employing L-Filters for Noise Reduction of Tomographic Data", Proceedings of the 2005 Finnish Signal Processing Symposium - FINSIG'05, 2005.
 - [23] J Mykkänen, J Tohka, J Luoma and U Ruotsalainen, "Automatic extraction of brain surface and mid-sagittal plane from PET images applying deformable models". Comput Methods Programs Biomed, vol. 79, pp. 1-17, 2005.
 - [24] Innis RB, Cunningham VJ, Delforge J, Fujita M, Gjedde A, Gunn RN, Holden J, Houle S, Huang SC, Ichise M, Iida H, Ito H, Kimura Y, Koeppe RA, Knudsen GM, Knuuti J, Lammertsma AA, Laruelle M, Logan J, Maguire RP, Mintun MA, Morris ED, Parsey R, Price JC, Slifstein M, Sossi V, Suhara T, Votaw JR, Wong DF, Carson RE (2007) Consensus nomenclature for in vivo imaging of reversibly binding radioligands. J Cereb Blood Flow Metab 27:1533–9, 2007.
 - [25] Gunn R. N., Lammertsma A. A., Humme S. P., Cunningham V.J., "Parametric imaging of ligand-receptor binding in PET using a simplified reference region model", NeuroImage 6, pp. 279-287, 1997.
 - [26] Wienhard, K.; Schmand, M.; Casey, M.E.; Baker, K.; Bao, J.; Eriksson, L.; Jones, W.F.; Knoess, C.; Lenox, M.; Lercher, M.; Luk, P.; Michel, C.; Reed, J.H.; Richerzhagen, N.; Treffert, J.; Vollmar, S.; Young, J.W.; Heiss, W.D.; Nutt, R., "The ECAT HRRT: performance and first clinical application of the new high resolution research tomograph", Nuclear Science Symposium Conference Record, Vol. 3, pp. 17/2 - 17/6, 2000.
 - [27] C. Xu, D.L. Pham, and J.L. Prince, "Medical Image Segmentation Using Deformable Models," Handbook of Medical Imaging -- Volume 2: Medical Image Processing and Analysis, pp. 129-174, edited by J.M. Fitzpatrick and M. Sonka, SPIE Press, May 2000.
 - [28] Jain, R., Kasturi, R., and Schunck, B. G., "Machine Vision", McGraw-Hill Inc, 1995.
 - [29] Yen, J.C., Chang, F.J., Chang, S., "A New Criterion for Automatic Multilevel Thresholding", IEEE Trans. on Image Processing, 4, pp. 370-378, 1995.
 - [30] Yeng, P.Y., Chen, L.H., "Random Sampling Thresholding: A New Approach to Multilevel Thresholding". Signal Processing, 34, pp. 311-322, 1993.
 - [31] Canny, J., "A Computational Approach to Edge Detection", IEEE Trans. on PAMI, 8, pp. 679-698, 1986.

- [32] Karaoui I., Fablet R., Boucher J.M., Augustin J.M., "Region-based image segmentation using texture statistics and level-set methods" ", IEEE international conference on Acoustics, Speech and Signal Processing, vol. 2, pp. 817-820, Toulouse, France, 2006.
- [33] McInerney T. and Terzopoulos D., "Topologically adaptable snakes", In Proc. Fifth International Conf. on Computer Vision (ICCV'95), Cambridge, MA, pp.840–845. Los Alamitos, CA: IEEE Computer Society Press, 1995.
- [34] M. Kass, A. Witkin, D. Terzopoulos, "Snakes: Active Contour Models", International Journal of Computer Vision, no.4, vol.1, pp. 321-331, 1988.
- [35] Li, Kang and Wu, Xiaodong and Chen, D. Z. and Sonka, M., "Optimal Surface Segmentation in Volumetric Images-A Graph-Theoretic Approach", Pattern Analysis and Machine Intelligence, IEEE Transactions on Vol. 28 No. 1, pp. 119-134, 2006.
- [36] http://en.wikipedia.org/wiki/Directed_graph#Directed_graph.
- [37] Xu Y.and Uberbacher E.C., "2-D Image Segmentation Using Minimum Spanning Trees," Image and Vision Computing, vol. 15, n. 1, pp. 47-57, 1997.
- [38] Falcão A.X., Udupa J.K., Miyazawa F.K., "An Ultra-Fast User- Steered Image Segmentation Paradigm: Live Wire on the Fly," IEEE Trans. Medical Imaging, vol. 19, Jan. 2000.
- [39] Wu Z. and Leahy R., "An Optimal Graph Theoretic Approach to Data Clustering: Theory and Its Application to Image Segmentation", IEEE Trans. Pattern Analysis Machine Intelligence, vol. 15, n. 11, pp. 1101-1113, Nov. 1993.
- [40] C. Alpert, A. Kahng, and S. Yao, "Spectral partition: The more eigenvectors, the better", Discrete Applied Math, 90, pp. 3-26, 1999.
- [41] J. Malik, S. Nelongie, T. Leung, and J. Shi, "Contour and texture analysis for image segmentation", Perceptual Organization for Artificial Vision Systems, Kluwer, 2000.
- [42] A. Y. Ng, M. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in Proc. of NIPS-14, 2001.
- [43] P. Chan, M. Schlag, and J. Zien, "Spectral k-way ratio cut partitioning," IEEE Trans. CAD-Integrated Circuits and Systems, vol. 13, pp. 1088–1096, 1994.

-
- [44] B. Kernighan and S. Lin, "An efficient heuristic procedure for partitioning graphs," *The Bell System Technical Journal*, vol. 49, no. 2, pp. 291–307, 1970.
 - [45] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 2000.
 - [46] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, pp. 1299–1319, 1998.
 - [47] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Math., Stat. and Prob.*, 1967, pp. 281–296.
 - [48] Inderjit S. Dhillon and Yuqiang Guan and Brian Kulis, "A Unified View of Kernel k-means, Spectral Clustering and Graph Cuts," *University of Texas Dept. of Computer Science*, 2005.
 - [49] Max-Planck-Institute for Neurological Research Cologne, Vinci Lite 2.37.0, 2001-2004.
 - [50] Liu Y., Collins R., Rothfus W., "Robust midsagittal plane extraction from normal and pathological 3-D neuroradiology images", *IEEE Transaction on Medical Imaging* 20(3), pp.175-192, 2001.
 - [51] The MathWorks, Inc., Matlab R2007a, 1984-2007.
 - [52] Kitware Inc., VolView 2.0 Free, 1996-2003.
 - [53] Sossi V, de Jong H, Barker W, Bloomfield P, Burbar Z, Camborde M, Comtat C, Eriksson L, Houle S, Keator D, Knöb, Krais R, Lammertsma A, Rahmim A, Sibomana M, Teräs M, Thompson C, Trebossen R, Votaw J, Walker M, Wienhard K, Wong D, "The second generation HRRT: a multi-centre scanner
 - [54] Comtat C, Bataille F, Michel C, Jones JP, Sibomana M, Janeiro L, Trebossen R, "OSEM-3D reconstruction strategies for the ECAT HRRT", *Conf Rec 2004 IEEE Nucl Sci Symp Med Imaging Conf (Rome 2004)*, Vol 6 3492–6, 2004.
 - [55] Jussi Hirvonen, Jarkko Johansson, Mika Teräs, Vesa Oikonen, Ville Lumme, Pauliina Virsu, Anne Roivainen, Kjell Någren, Christer Halldin, Lars Farde and Jarmo Hietala, "Measurement of striatal and extrastriatal dopamine transporter binding with high-

- resolution PET and [11C]PE2I: quantitative modeling and test–retest reproducibility”, *Journal of Cerebral Blood Flow & Metabolism*, pp. 1-11, 2008
- [56] Diana Martinez, Mark Slifstein, Allegra Broft, Osama Mawlawi, Dah-Ren Hwang, Yiyun Huang, Thomas Cooper, Lawrence Kegeles, Eric Zarahn, Anissa Abi-Dargham, Suzanne N. Haber, and Marc Laruelle, “Imaging Human Mesolimbic Dopamine Transmission With Positron Emission Tomography. Part II: Amphetamine-Induced Dopamine Release in the Functional Subdivisions of the Striatum”, *Journal of Cerebral Blood Flow & Metabolism*, vol. 23, pp.285–300, 2003.
- [57] Duda, R.O., Hart, P.E., Stork, D.G., “Pattern Classification”, 2nd edition, John Wiley & Sons, Inc., 2001.

APPENDIX 1: Eigenvectors and Eigenvalues

The information given in this appendix is taken from [57].

Given a d -by- d matrix M , a very important class of linear equations is of the form

$$Mx = \lambda x \quad (19)$$

for scalar λ , which can be rewritten

$$(M - \lambda I)x = 0 \quad (20)$$

where I is the identity matrix and 0 is a zero vector. The solution vector $x = e_i$ and the corresponding scalar $\lambda = \lambda_i$ are called the eigenvector and associated eigenvalue, respectively. If M is real and symmetric, there are d (possible nondistinct) solution vectors $\{e_1, e_2, \dots, e_d\}$, each with an associated eigenvalue $\{\lambda_1, \lambda_2, \dots, \lambda_d\}$. Under multiplication by M the eigenvectors are changed only in magnitude, not direction:

$$Me_j = \lambda_j e_j \quad (21)$$

If M is diagonal, then the eigenvectors are parallel to the coordinate axes.

One method of finding the eigenvectors and eigenvalues is to solve the characteristic equation (or secular equation,)

$$|M - \lambda I| = \lambda^d + a_1 \lambda^{d-1} + \dots + a_{d-1} \lambda + a_d = 0 \quad (22)$$

for each of its d (possible nondistinct) roots λ_j . For each such root, one can solve a set of linear equations to find its associated eigenvector e_j .

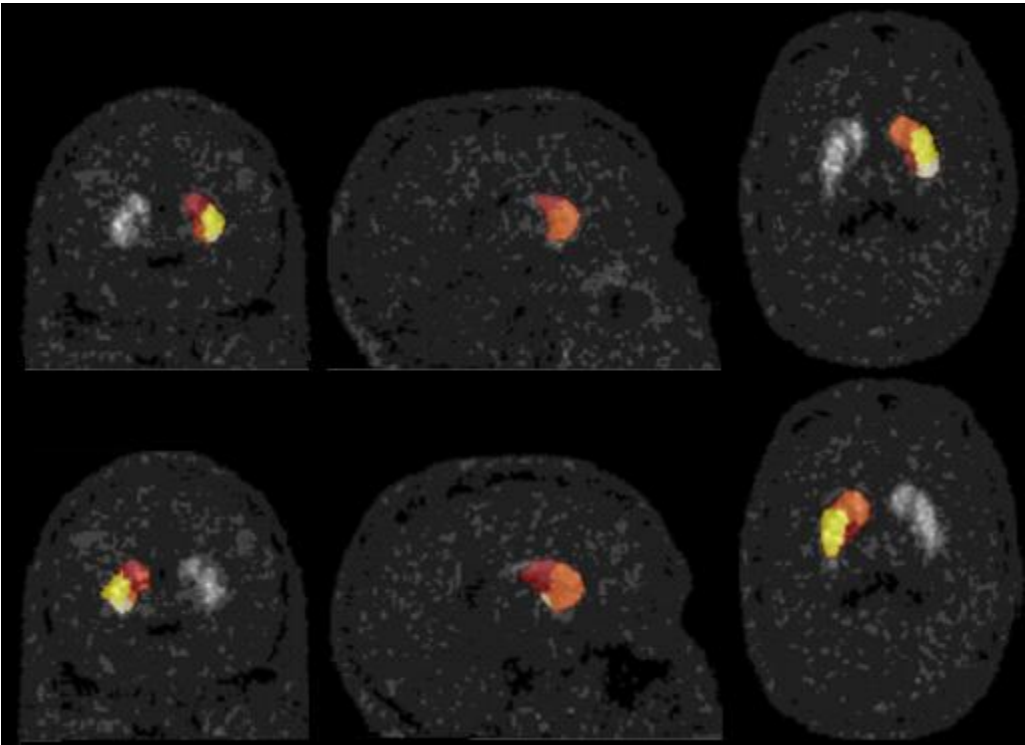
Finally, it can be shown that the trace of a matrix is just the sum of the eigenvalues and the determinant of a matrix is just the product of its eigenvalues:

$$\text{tr}[M] = \sum_{i=1}^d \lambda_i \text{ and } |M| = \prod_{i=1}^d \lambda_i \quad (23)$$

If a matrix is diagonal, then its eigenvalues are simply the nonzero entries on the diagonal, and the eigenvectors are the unit vectors parallel to the coordinate axis.

APPENDIX 2: Segmentation results

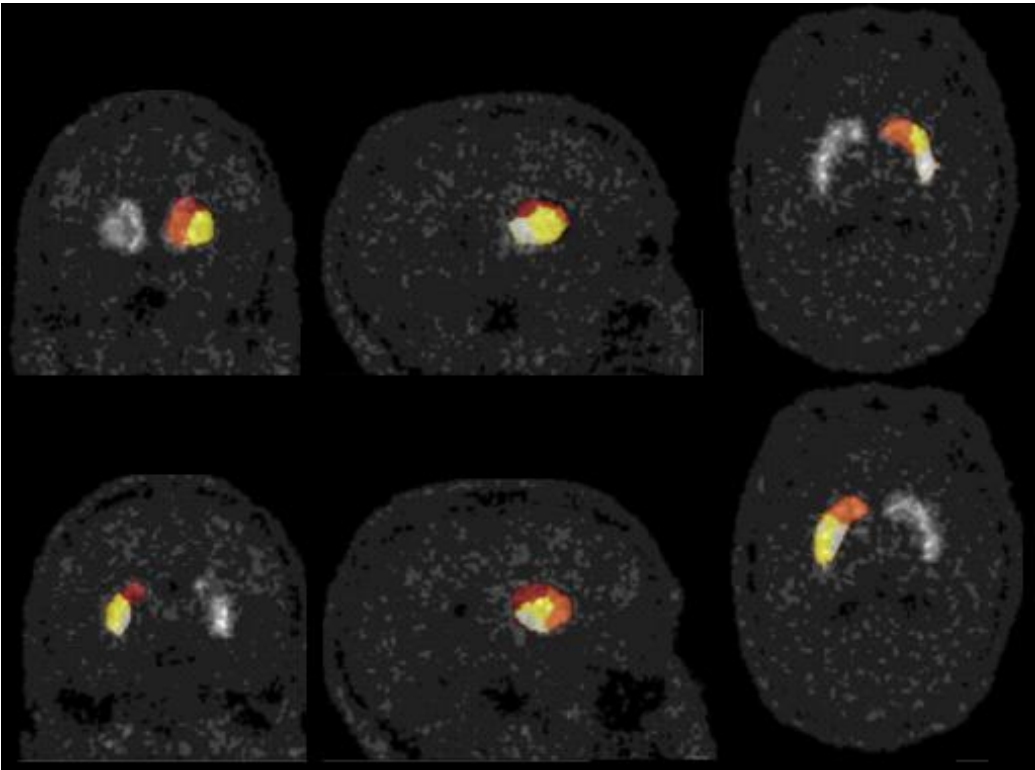
Image1 (name: rb0008bp):



Slice: 94

Slice: 92 / 116

Slice: 81

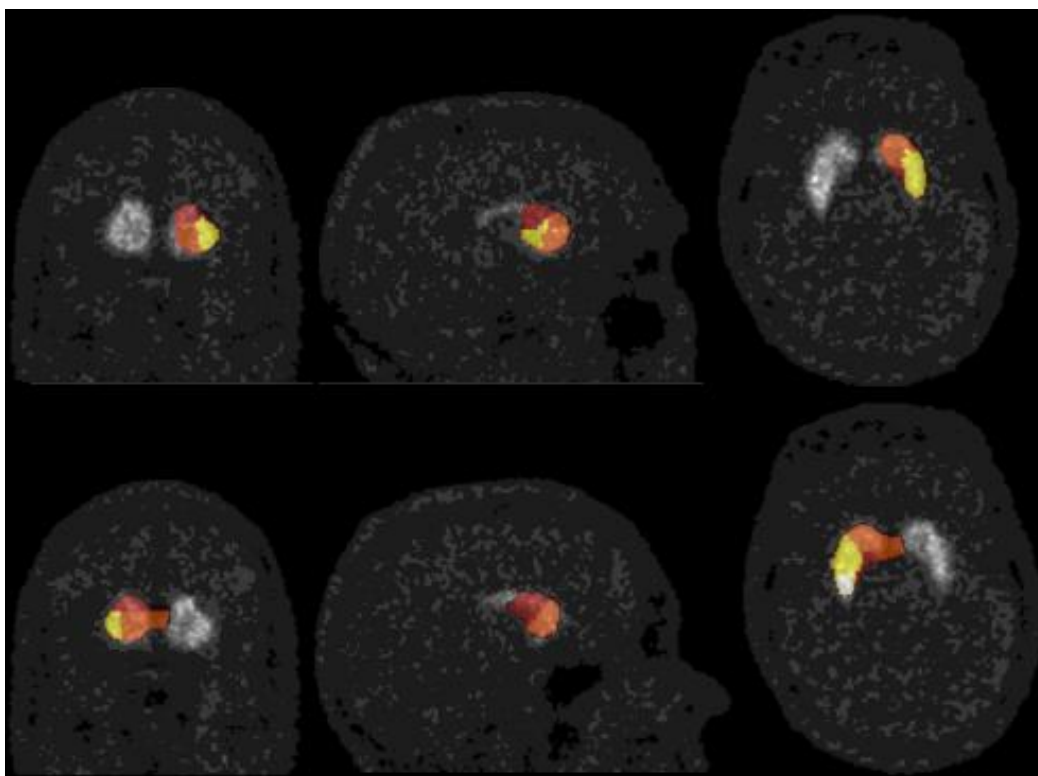


Slice: 88 / 102

Slice: 82 / 120

Slice: 88

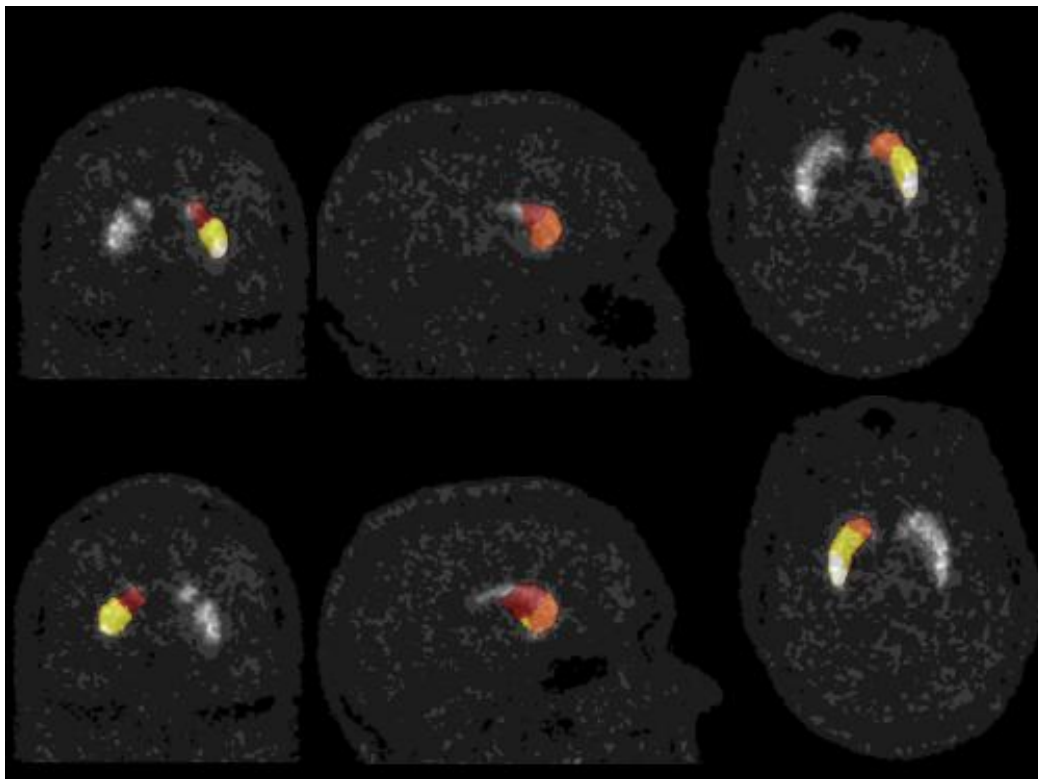
Image2 (name: rb0009bp):



Slice: 95

Slice: 92 / 116

Slice: 81

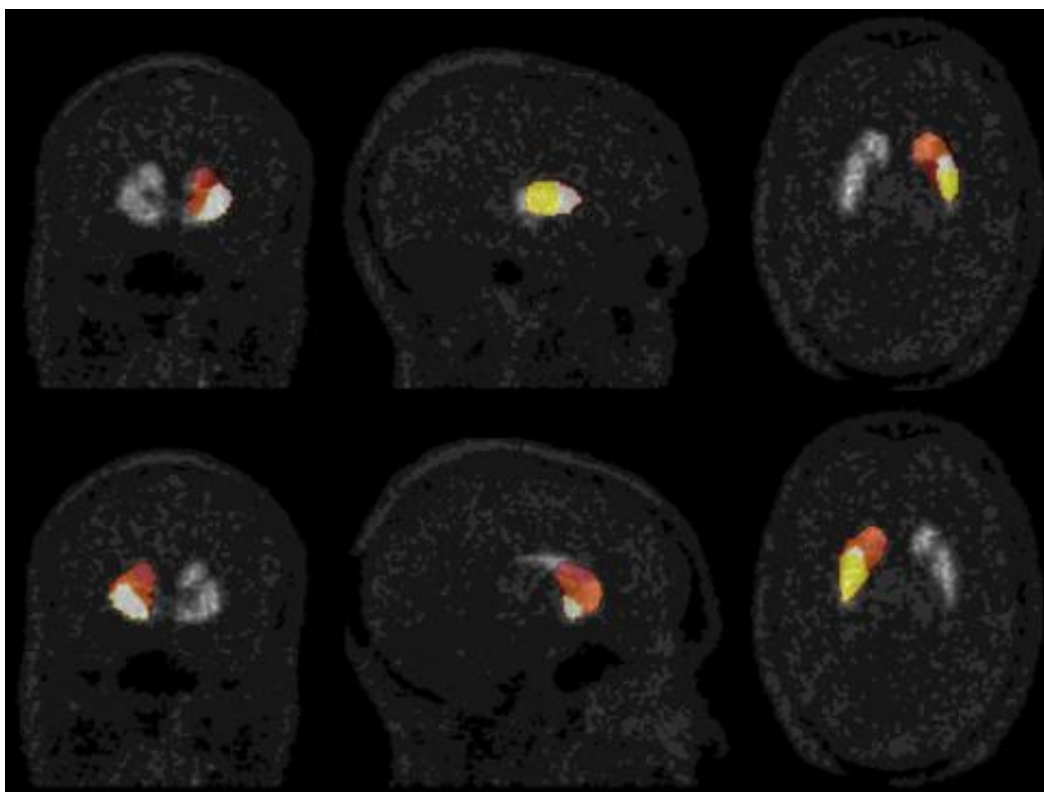


Slice: 105

Slice: 98 / 120

Slice: 88

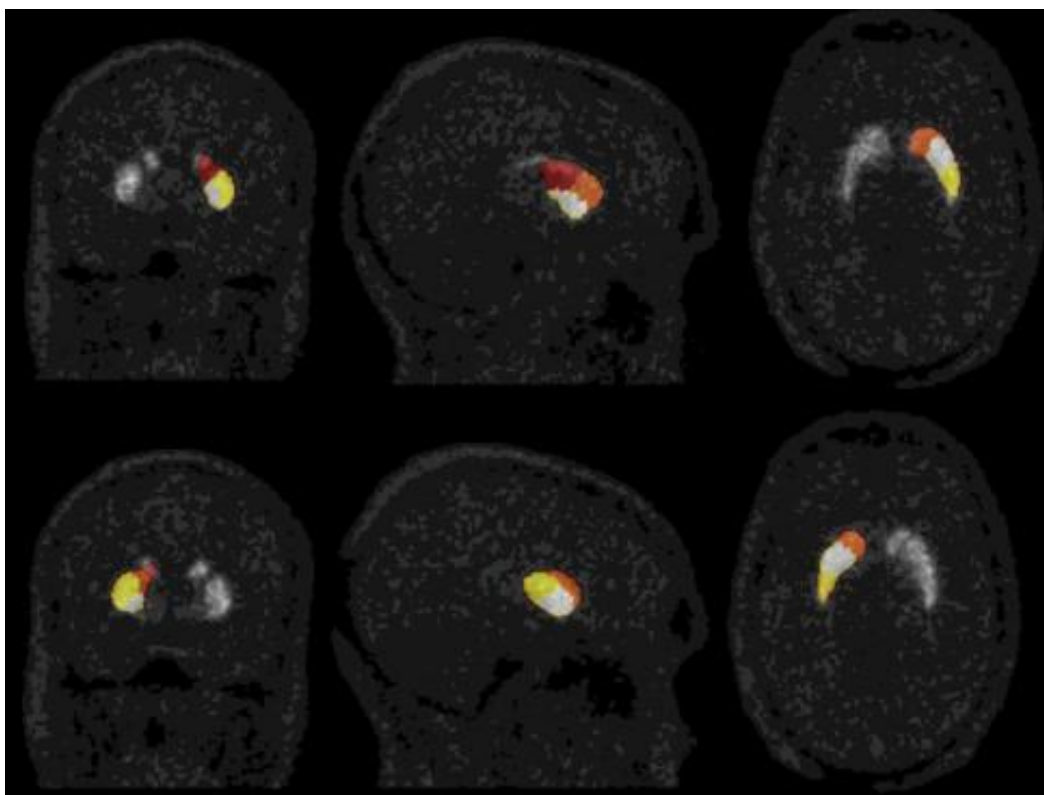
Image3 (name: rb0031bp):



Slice: 108

Slice: 106 / 140

Slice: 113

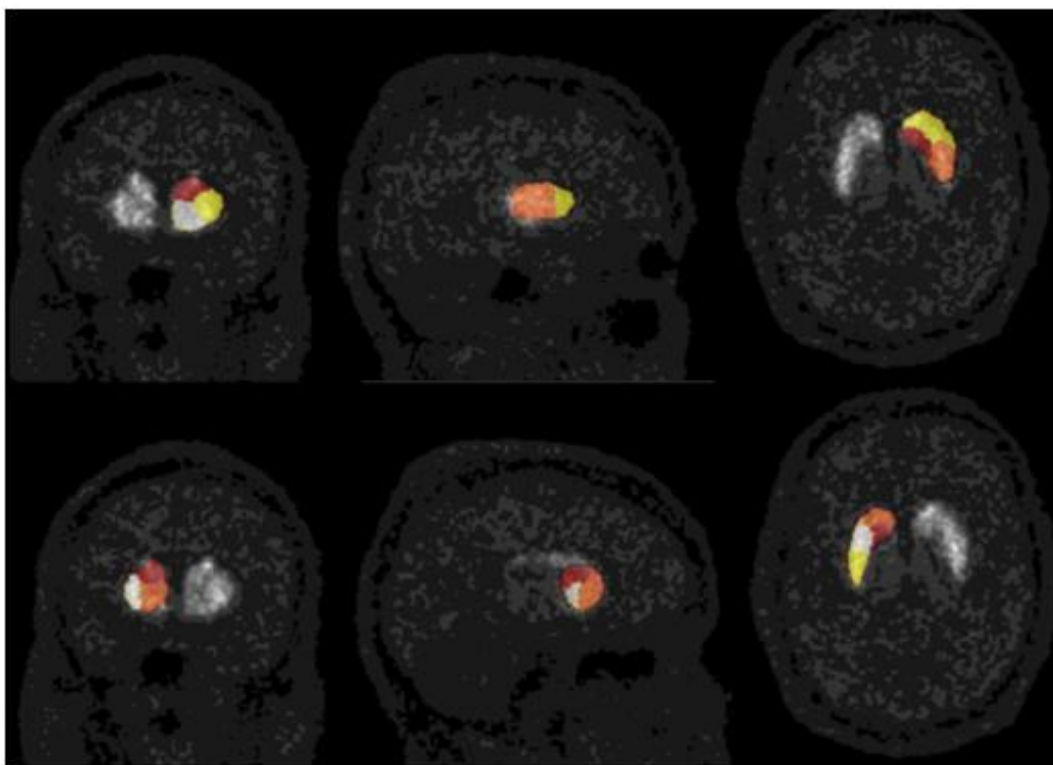


Slice: 116

Slice: 114 / 146

Slice: 122

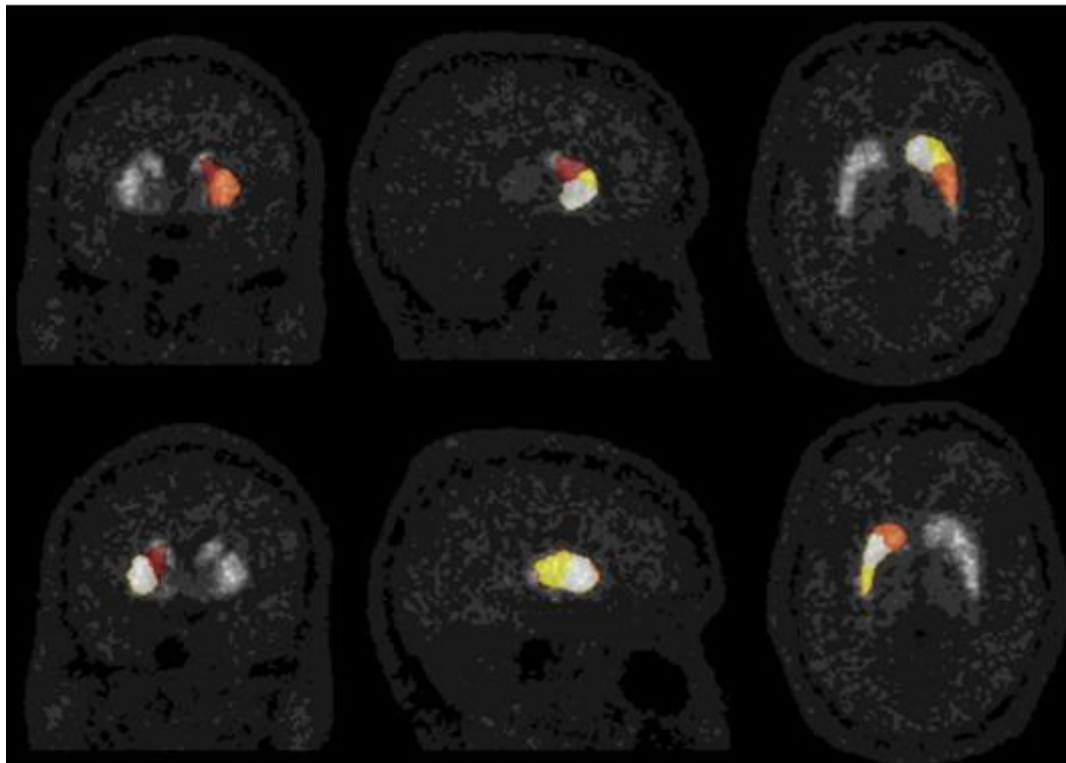
Image4 (name: rb0037bp):



Slice: 83

Slice: 80 / 116

Slice: 69



Slice: 90/88

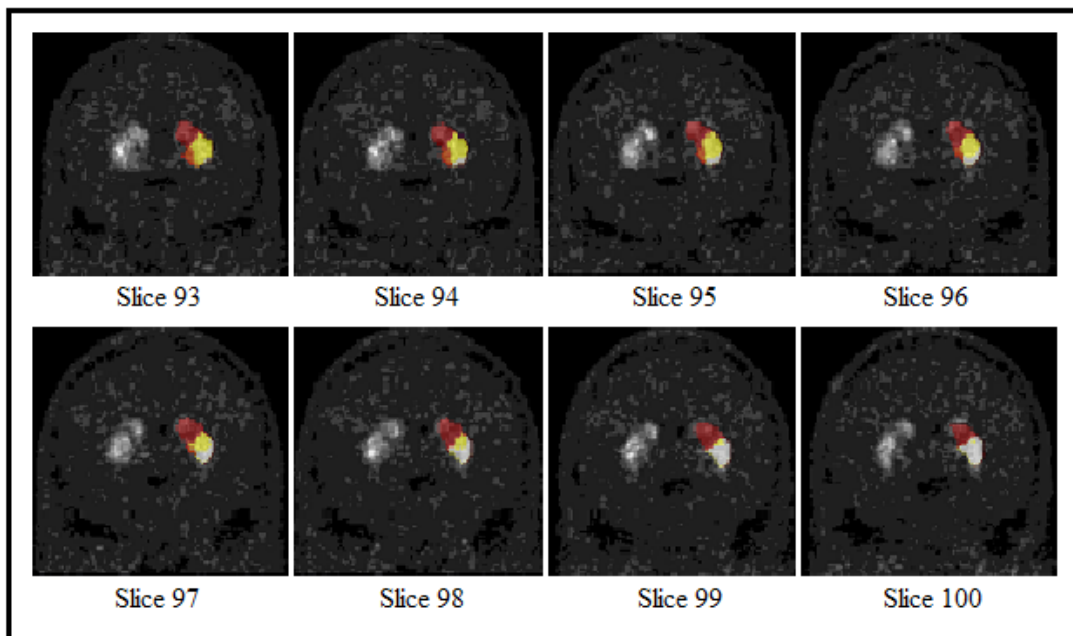
Slice: 92 / 123

Slice: 77

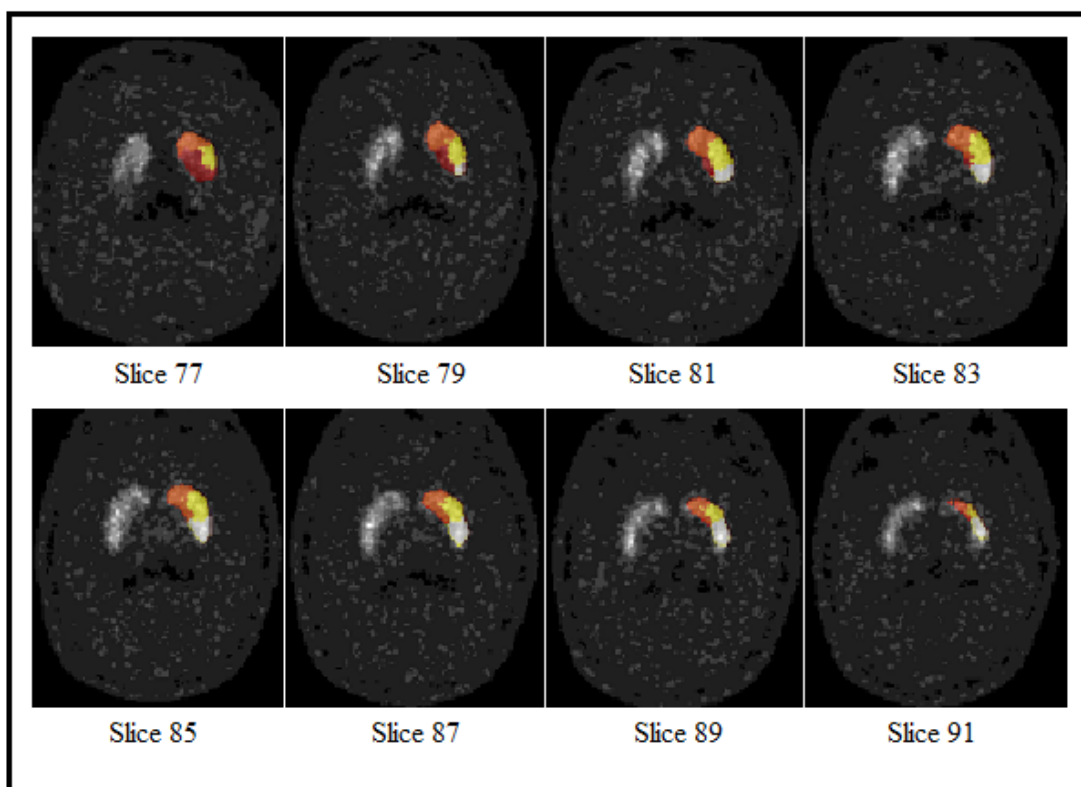
APPENDIX 3: Putamen subdivision

Image1 (name: rb0008bp):

Coronal view:



Transverse view:



Sagittal view:

