

UNIVERSITAT POLITÈCNICA DE CATALUNYA  
DEPARTAMENT DE LLENGUATGES I SISTEMES INFORMÀTICS  
**MÀSTER EN COMPUTACIÓ**

## TESI DE MÀSTER

### **Automatic Computation of Potential Tumor Regions in Cancer Detection using Fractal analysis techniques**

ESTUDIANT: Bernat Orellana Bech  
DIRECTOR: Pere Brunet i Crosa

DATA: 8 DE SETEMBRE DE 2009.



## ABSTRACT

The initial goal of this master thesis has been to explore the capabilities of Fractal Analysis and its applications to tumor detection on volumetric medical images. Study cases include images obtained with Computer Tomography or Magnetic Resonance equipment.

This document refers to the current state-of-the-art of fractal analysis, and its applications to medicine. More in particular, it focuses on the detection, segmentation and classification of tumors in volumetric images. Fractals have proved to be a useful tool for image analysis. In this field, the study of multifractal spectrum, based on the multifractal formalism, has had a remarkable importance. In addition, a new approach called isofractal spectrum is proposed, which relies on the boxcounting fractal dimension calculation of iso-density surfaces.

Based on these fractal techniques, a software module has been developed with the purpose of easing the detection of tumoral tissues found in MRI/CT liver images. The work has been divided in two big stages, the first one focuses on the construction of the fractal analysis tools; the second deals with the study of tumoral and healthy areas and its fractal spectrum. This process has allowed defining spectrum descriptors, which are the core of tissue classifiers. Finally, the processes are applied to the data sets and the classification results are reviewed.

# Table of Contents

1	Introduction and Motivation.....	1
2	Organization .....	2
3	Techniques for medical image acquisition .....	3
3.1	Radiography .....	3
3.2	Computed Tomography.....	4
3.3	Magnetic Resonance Imaging (MRI) .....	5
3.4	General purpose parameters of CT and MRI volume data.....	8
3.5	Positron Emission Tomography .....	8
4	Fractals .....	10
4.1	The Iterated Function System .....	10
4.2	Classic Fractals and Self-Similarity .....	11
4.2.1	Cantor set.....	12
4.3	Deterministic and Natural fractals.....	12
4.3.1	Euclidean and Topological Dimension .....	13
4.3.2	Self-similar dimension .....	13
4.3.3	Compass-Dimension .....	14
4.4	Multifractal theoretical overview .....	17
4.4.1	Introduction .....	17
4.4.2	Definitions .....	17
4.4.3	Methods for multifractal spectrum calculation. ....	18
4.4.4	Basic properties of the multifractal spectrum.....	21
5	Previous work.....	22
5.1	Fractals and Cancer. Is there a relationship?.....	22
5.1.1	Angiogenesis and statistical growth processes.....	22
5.1.2	Imaging using contrast agents .....	24
5.1.3	Fractals and Cancer .....	24
5.2	FA: Fractal Analysis on medical images.....	24
5.2.1	Study cases .....	24
5.3	MFA: Multifractals Analysis .....	27

5.3.1	Multifractal Analysis: an interdisciplinary tool.....	27
5.3.2	Multifractals applied to cancer detection and segmentation. ....	27
6	Overview of the Application for hepatic Tumoral tissue detection.....	33
6.1	System Overview .....	34
6.2	Data sets .....	35
7	Definition and Implementation of Fractal Analysis Algorithms for Tumor detection .....	36
7.1	Isofractal spectrum analysis .....	36
7.1.1	Isosurface Box counting.....	36
7.1.2	Isofractal spectrum .....	38
7.1.3	Problems with resolution in volumetric models.....	42
7.2	Multifractal spectrum analysis .....	42
7.2.1	Validation .....	44
7.2.2	The MF spectrum scanline .....	44
7.3	Application of the fractal analysis techniques.....	45
7.3.1	Global operations .....	45
7.3.2	Local operations .....	46
7.3.3	Dynamic operations.....	47
7.4	Summary .....	47
8	Analysis and Classification of MRI based on fractal characterization .....	48
8.1	Methodology .....	48
8.2	Characterization phase .....	49
8.2.1	In-and-out scanlines .....	49
8.2.2	Multiple area comparison.....	51
8.3	Second phase: Descriptors definition.....	54
8.3.1	Isofractal spectrum descriptors.....	54
8.3.2	Multifractal spectrum descriptors.....	54
8.4	Third Phase: Classifier definition.....	55
8.4.1	Static classifiers.....	55
8.4.2	Dynamic classifiers .....	56
9	MRI Classifiers and Results .....	57
9.1	Final classifiers.....	57

9.2	Region voting and visualization .....	58
9.3	MRI Classification Results.....	58
10	Analysis and Classification of CT based on fractal characterization .....	62
10.1	Introduction .....	62
10.1.1	Fractal analysis characterization .....	63
11	Conclusions and Future work .....	65
12	Bibliography .....	67
13	Appendices .....	69
13.1	Synthetic Test cases for multifractal and isofractal spectrum. ....	69
13.2	Medical image sample cases for multifractal and isofractal spectrum. ....	75
13.2.1	Homogeneous liver area ( $l=32, \{118,238,37\}$ ).....	76
13.2.2	Homogeneous Tumoral border area ( $l=32, \{113,220,47\}$ ) .....	77
13.2.3	Non-Homogeneous Tumoral border area ( $l=32, \{152,202,47\}$ ) .....	78
13.2.4	Inside the Tumoral area ( $l=32, \{128,213,47\}$ ) .....	79
13.2.5	Liver border and vessels ( $l=32, \{173,236,38\}$ ) .....	80
13.2.6	Scan-line in and out the metastasis. 64 samples path. ....	81
13.3	User Interface .....	82
13.3.1	Static selector .....	82
13.4	Dynamic selector.....	83
13.5	Local selection.....	84
13.6	Test cases .....	84
13.7	DimFract module Class Diagram .....	86

# 1 INTRODUCTION AND MOTIVATION

Radiology is one of the most active and technologically advanced fields in medicine. It was born from the most advanced physics concepts, and it became a reality thanks to the state-of-the-art of electronics and computer science. The advances of medical imaging have made possible the early detection and diagnosis of multiple affections that were not at our reach just some years ago. However, progress comes with a price. The raise of imaging machinery has implied that the number and complexity of technical parameters have grown in the same proportion, and the amount of information generated by the imaging devices is much larger.

In spite of technical progress, medical imaging supply chain invariantly finalizes at the same point: a human being, typically the radiologist or medical practitioner in charge to interpret the obtained images. At the end, it is not unusual that human operators check one by one two hundred slices of a Computer Tomography coming from a single routine control scanner. It is not surprising if some tiny detail is missed when searching for “something wrong”, especially after some hours of continuous visualization, or due to insufficient time budgets. One of the milestones of this work is providing the reader with an overview of the field of volumetric medical imaging, in order to achieve a sufficient understanding of the problematic involving this discipline.

This master thesis is mainly an exercise of exploration of a set of techniques, based on fractal analysis, aimed to provide any sort of computational help to the personal in charge of the interpretation of volumetric medical images. Fractal analysis is a set of powerful tools which have been applied successfully in multiple fields. The thesis goal has been to apply these techniques within the scope of tumor detection on liver tissues and evaluate their efficiency and adequateness.

## 2 ORGANIZATION

This document is organized in three parts:

- First part introduces the theoretical background required for the document. Medical image acquisition techniques are explored in Section 3. Section 4 covers the Fractals theory, as well as the Multifractal theoretical basis.
- Second part, covered in Section 5, describes previous work related to the application of Fractal analysis in medical imaging, and more concretely on cancer detection.
- Third part is devoted to the study of developments performed in the master thesis, aimed to build a tumor detector in MRI liver images. The contents can be found distributed in the chapter 6 to 11. Chapter 6 offers an encompassing view of the tumor detection system and its environment. It follows the description of the fractal analysis algorithms used for tissue analysis and their implementation (chapter 7). The next step is the characterization of malignant areas and the construction of a classifier, covered in chapter 8. The final classification results are shown in chapter 9.

Chapter 10 covers some of the preliminary work done with CT images, offered just as a complementary view of the MRI imaging.

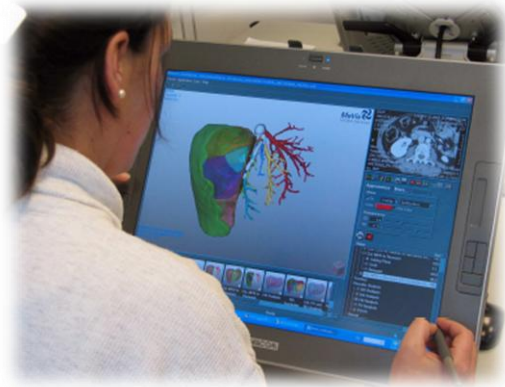
Finally, conclusions and future work are detailed in section 11.



### 3 TECHNIQUES FOR MEDICAL IMAGE ACQUISITION

Imaging techniques have had a significant influence in the development of modern medicine. They cover a wide range of applications, from educational purposes, as a teaching tool for future practitioners, to diagnosis and treatment planning. The information obtained by the means of medical imaging may be crucial for granting an early detection and an accurate diagnosis and staging without the need of applying invasive techniques.

Diagnosis can benefit from 2D and 3D interactive visualization of acquired data. The imaging systems do not limit themselves to display a raw image obtained from medical equipment, but image processing algorithms are applied, allowing to extract hidden data from images, which otherwise would remain invisible to the human eye. Images are employed to assess multiple organic parameters such as tissue density, blood perfusion, contrast agent enhancement, or metabolic behavior.



**Figure 3.1: Liver Surgery Planning**

Lately, the support of imaging has been extended to interoperative support, the medical visualization systems have entered inside the operating room. Preoperatively acquired images and intraoperative images are integrated to provide support during the operation. Surgeon is

supported by Virtual Reality or mixed reality devices, such as see-through displays, semi-transparent head-mounted displays or simple projectors, which provide instant information of the data required.



**Figure 3.2: Intraoperative imaging projection in liver surgery**

Another growing area is the surgery simulation trainers, which combines visualization and haptic devices in order to permit the surgeon to have a hands-on trial before the real thing. Arthroscopic knee/shoulder surgical simulators<sup>1</sup> or cataracts training simulation [2] are interesting examples of that area.

For further information, an overview of the state-of-the art of imaging on medicine could be found at [1].

#### 3.1 Radiography

The discovery of X-Rays in 1895 by Wilhelm Konrad Rontgen was a revolutionary event in the history of medicine. For the first time it was possible to see interior parts of the body without actually cutting into it.

X-ray imaging is based on the attenuation of X-ray quanta when travelling through the body. The remaining quanta reach a film, where a 2D image is formed. From the physical point of view, X-ray attenuation is caused by two different sources, the absorption and the scattering (or

<sup>1</sup> [http://www.insightarthrovtr.com/index\\_en.htm](http://www.insightarthrovtr.com/index_en.htm)

Compton effect). Absorption is the natural capture of the quanta within the tissue, which is the predominant effect for low energy X-ray beams (so called soft beams). In the other hand, scattering is provoked by the collision of high energy beams (hard beams) with electrons from the outer hull of an atom.



**Figure 3.3: Wilhelm Conrad Röntgen wife's hand. The first X-ray photograph.**

X-rays are generated by electrons accelerated through an electric field. When the electrons hit the anode, the kinetic energy is transformed into heat and X-ray quanta. The speed of electrons depends on the voltage applied during the field generation. We speak of hard beams when the voltage goes beyond 100kV.

Scattering is the dominant attenuation effect for very dense tissues. Hence, hard beams are suitable in the acquisition of bones structures. When the goal is soft-tissue exploration, which is the case of mammographies, soft beams are applied. Since the dominant attenuation process for soft beams is absorption, it implies a higher impact of the radiation on the tissues, which makes it more harmful than hard beams.

The images obtained using X-ray are a planar representation of the attenuation provoked by the whole pass of radiation through the body in a given direction. Hence, there is a natural integration of the attenuation provoked for different overlapped tissues found in the path of a ray. The result is that we obtain a 2D representation

from a 3D object, a kind of projection, where it is difficult to locate precisely the organic structures, or its composition. In the other hand, those structures having its silhouettes orthogonal to beam's direction are relatively well defined. In consequence, it is usual to take several images from different viewpoints, in order to capture details of the regions of interest. Imaging with clarity inner tissues or organs hidden behind other body structures are out of the reach of this technique.

## 3.2 Computed Tomography

Computed tomography also known as CT<sup>2</sup> was the first device capable of providing a volumetric representation of an object. CT scanner obtains a real 3D representation of the target object, which is formed by an array of 2D images representing slices of the scanned object.

Just as radiographies, CT are based on the attenuation of X-ray crossing the human body, but in this case the process for obtaining each of the slice-images it is quite more sophisticated. The scanning device has an X-ray source and an array of sensors attached to a rotating ring, each slice is the result of the intensity received by the detectors after a full loop. Intensity signal is post-processed using the Radon inverse transform [3], in order to compose a unique image. CT sensors read the integration of the X-ray attenuation of multiple beams, which coincides with the Radon transform mathematical definition. The problem here is to recover the model from the integration data, which is performed by the inverse Radon transformation.

After an image has been calculated, the device is positioned for the next slice just by performing an axial step.

<sup>2</sup> The Spanish acronym is TAC: "Tomografía Axial Computarizada"

Notwithstanding the higher amount of radiation exposition required for CT scanners, which may be especially risky for young children, the advantages of CT over radiographies are overwhelming:

- CT allows locating and obtaining a representation of inner anatomical structures. While X-ray images integrate X-ray absorption of multiple tissues crossed by the beams, CT records the attenuation provoked for small volume elements named *voxels*.
- The sensitivity of CT is about two orders of magnitude greater than X-ray images. X-ray images are not able to distinguish different soft tissues, whereas CT images permit a primary segmentation. It must be remarked that CT resolution may not suffice for certain applications involving soft tissues, which will require a different approximation: the MRI scanner.
- CT is able to obtain high precision measures of the X-ray absorption for each voxel. This quantitative measurement of CT voxels is normalized and expressed in Hounsfield units (*HUs*). Precise and deterministic measurement is useful when determining the status of a given tissue, for instance, *HUs* help determining the bone density in osteoporosis diagnosis. It is also an important property for computerized aided diagnostic, since it is easier for engineers to establish invariant relationships between certain ranges of HUs and tissue types.

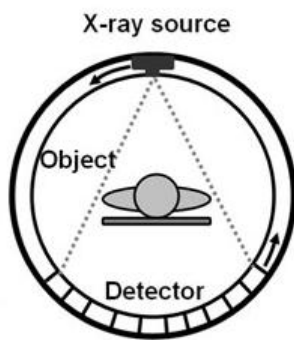


Figure 3.4: CT scanner schema



Figure 3.5: CT scanner and control room

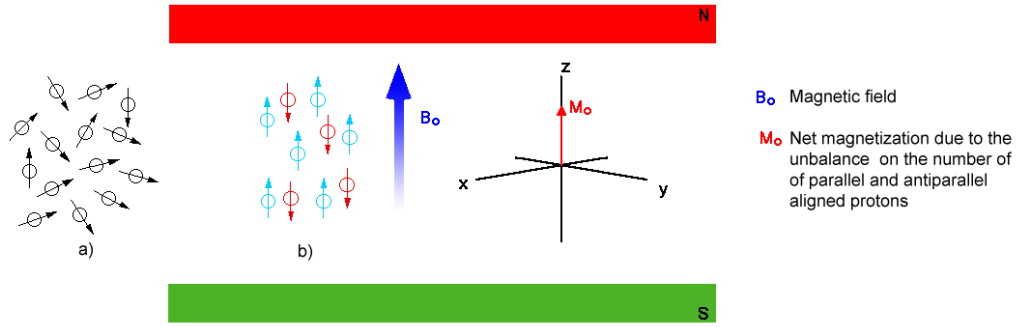
### 3.3 Magnetic Resonance Imaging (MRI)

MRI is an imaging technique capable of obtaining a volumetric representation of the body. The image intensity values are related to the density of water present on the tissues. MRI applications are rather complementary to CT, MRI is adequate for soft-tissues classification due to their high water content, but its performance is poorer in other areas such as skeletal structures, where CT excels. MRI provides a high soft tissue contrast, being the perfect choice for neuro-imaging and joint diagnosis (shoulders and knees). Moreover, because of it relies in magnetism and do not uses ionizing radiation, it is not harmful for the patient. One of the main drawbacks for MRI is the complexity of interpretation of the resulting images, which usually require the participation of a specialized radiologist in addition to the physician. The difficulty is related to the large set of parameters that involve MRI captures. The final and most limiting constraint which impedes a further extension of MR imaging is the cost, which explains why CT equipment are found more frequently in the hospitals than MRI.

The physic foundations of magnetic resonance are deeply related with the quantum mechanical property of spin, which characterizes subatomic particles such as protons. Certain nuclei such as  $^1\text{H}$  (protons),  $^2\text{H}$ ,  $^3\text{He}$ ,  $^{23}\text{Na}$  or  $^{31}\text{P}$ , have a non-zero spin, which provokes a magnetic moment. For organic tissues analysis, the most advantageous nucleid is the hydrogen isotope  $^1\text{H}$ , since it

is the most abundant of the three existing hydrogen isotopes (99,85%) , and its biological abundance in the human body is about 63%, which is basically related to the presence of water within the tissues.

$^1\text{H}$  atom is composed by a single proton, which has two spin states: "up" and "down". When protons are placed into a strong external magnetic field, their spins change, aligning with the direction of the field. Protons fall in one of the two available energy states, which are separated by a certain splitting energy. The total sum of individual proton's magnetic moments is known as Net magnetization (M).



**Figure 3.6: a) natural state of proton spins. b) effect of the MRI magnetic field on the H protons spin alignment, and the resulting net magnetization.**

Once the protons are spin-aligned, it is possible to change the Net magnetization by exposing them to a RF (radio frequency) pulse able to deliver the energy difference between the spin states. The frequency of the pulse must be a certain value known as Lamor frequency. If enough energy is put into the system, it is possible to saturate the spin system and make  $M_z=0$ . The time which describes how  $M_z$  returns to its equilibrium after the pulse is terminated is called the spin-lattice relaxation time or T1.

Immediately after the pulse is emitted, all protons spin turn on a similar phase. After few milliseconds the phases of the different protons recover its random distribution. The time to reach that state is called the spin-spin relaxation time, T2.

Both T2 and T1 parameters are measured and combined for determining the  $^1\text{H}$  proton density ( $\sigma$ ) present in a volume element (or voxel), and hence, the water presence. For instance, cerebrospinal fluid has 97% water content, while gray and white matter has respectively an 84 % and 71%.

$$\sigma = k \rho \left(1 - e^{-\frac{T_R}{T_1}}\right) \cdot e^{-\frac{T_E}{T_2}}$$

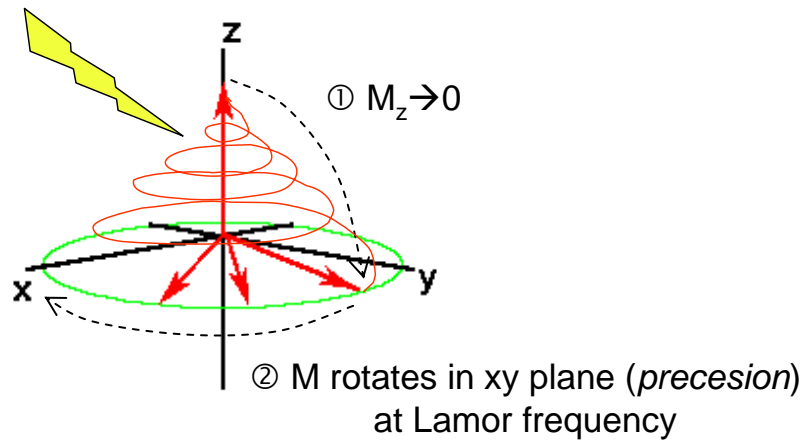
The influence of T1 and T2 effects on  $\sigma$  calculation depends on the times between the initial stimulation and the measurements (echo time  $T_E$ ) and the time between two stimulation cycles (repetition time  $T_R$ ). Adjusting those parameters, the operator may select the weight of T1 or T2 phenomena in the measurement, which leads to three cases:

- T1-weighted images (short  $T_R$  and  $T_E$ ).
- T2-weighted images (long  $T_R$  and  $T_E$ ).
- Proton-weighted (long  $T_R$  and short  $T_E$ ).



**Figure 3.7: Spinal Cord MRI : Left T1-weighted. Right T2-weighted**

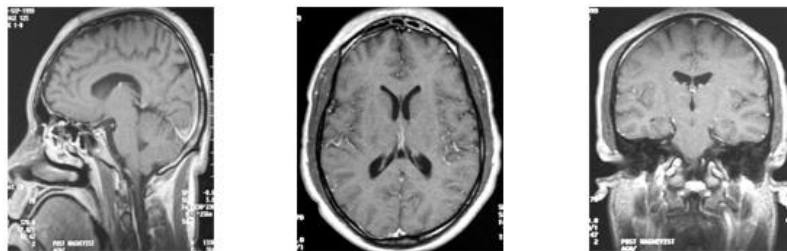
① RF photon stimulus



**Figure 3.8: Effects of RF pulse at Larmor frequency on the magnetic moment of a single proton.**

In order to reconstruct the spatial information of the signal, two additional gradient magnetic fields are required, one enables the selection of the slice in  $z$  direction, while the other selects a  $y$  row within the current slice.

One important advantage of MRI is that the slice direction can be configured at will, which is not possible with CT scanners, where the slice direction is always axial. This is convenient in order to ease medical analysis of the images.



a) Saggital scan

b) Axial scan

c) Coronal scan

**Figure 3.9: Several slice directions are available using MRI**

MR scanner is performed in a ring magnet that is large enough to enclose the whole patient. Inside this magnet, gradient coils are embedded in such a way that magnetic field gradients in three orthogonal directions ( $x$ ,  $y$  and  $z$ ) can be generated.

For further information on physics and mathematical of MRI, please refer to [4].



One of the downsides of MRI versus CT is that there is not a standard density scale. This means that the intensities in MR images do not have a fixed tissue-specific numeric meaning, not even within the same MRI protocol, for the same body region or even for the images of the same patient obtained on the same scanner.

### 3.4 General purpose parameters of CT and MRI volume data

Common parameters of datasets obtained using MRI and CT are related to the size and spatial resolution of the images. Basically, we should take care of:

- **Number of slices:** it is the total number of image planes that form the whole dataset. It may vary depending on the scanning range (whole body, abdomen, head...) and the axial resolution of the scanner.
- **Number of pixels per slice:** Refers to the image matrix size, measured in voxels x voxels. Typical size is 512x512
- **Intraslice resolution (Voxel size within a slice):** Represents the real distance of the slice-adjacent voxels in the scanned body.
- **Interslice resolution (Voxel size along axial direction):** It represents the gap between a slice and the following.

CT and MRI provide a comparable spatial resolution, but MRI offers a better sensitivity or contrast resolution. For all the cases, interslice resolution is always lower than intraslice resolution; hence, we speak of anisotropic datasets. Typical resolutions are in the range of millimeters.

High definition CTs are available, which provide a smaller intraslice, at the cost of a higher exposition to radiation. Given its physics foundations, MRI allows more isotropic data fields, granting that the intraslice distance will be closer to the in-train image distance.

### 3.5 Positron Emission Tomography

FD-PET is a nuclear medicine volumetric imaging technique which permits displaying functional processes of the body.

The tracer, a short-lived radiopharmaceutical substance, is injected and tracked throughout the body till it reaches its target area. Once there, the substance is processed by the metabolism and

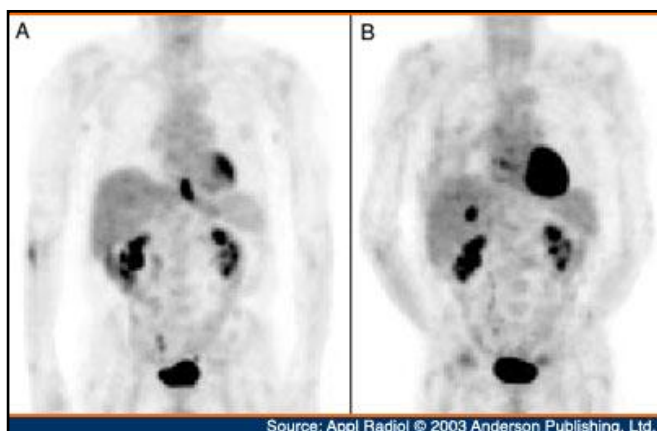


Figure 3.10: FDG-PET Scan

the radioactive isotope decays, emitting positrons, which in turn may collide with electrons, finally generating two gamma photons in opposed directions. Those photons are captured by sensors, and that provide information of the 3D coordinates. This technique allows a volumetric reconstruction of metabolism information, providing dynamical information of the body mechanisms.

In modern scanners, the reconstruction is often accomplished with the aid of a CT X-ray scan, which is performed simultaneously

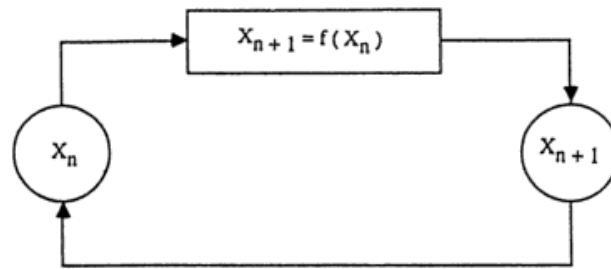
during the same session, in the same machine. Resolution of PET is much lower than CT (2-3mm per pixel), but the information provided can be precious.

If the biologically active molecule chosen for PET is of fluoro deoxyglucose, similar to glucose, the concentrations of tracer tend to occupy areas with a high metabolic activity in terms of regional glucose uptake. This property is of the maximum interest in tumor detection and aggressiveness classification. Standardized uptake value, SUV, is a PET quantifier calculated as a ratio of tissue radioactivity concentration at a time  $T$ ,  $C_{PET}(T)$ , the injected dose and the body weight. Cancer treatment response is usually assessed with FDG PET by calculating the SUV on the tissues affected by tumors.

## 4 FRACTALS

### 4.1 The Iterated Function System

Iterated Function Systems (IFS) are the foundations of fractals. The ingredients for the simplest IFS are: an initial input set  $x_0$ , a transformation function  $f(x)$ , and a simple iteration process. For each iteration step, the system will apply the function to the input set, creating a new output set. At each iteration, except the first one, the input of  $f(x)$  is the output of the previous iteration, hence we speak of a feedback process, defined by the equation  $x_{n+1} = f(x_n)$ . Iterating the process, the output sets  $x_1, x_2, \dots, x_n$  are generated in the steps  $1, 2, \dots, n$ . It is understood that the result of an IFS is the result of infinite iterations, i.e.  $\lim_{n \rightarrow \infty} x_n$



**Figure 4.1: A simple Iterative Function System schema**

In order a system is to be considered a IFS, an extra condition must be fulfilled: it is required that transformation function  $f(x)$  accomplishes the contraction mapping principle<sup>3</sup>, which intuitively could be interpreted as a condition enforcing the function to cause reductive or contractive space transformations on the original set. In other words,  $f(x)$  set should be a spatial contraction of the original set  $x$ . If this condition is accomplished, despite of the initial value of  $x_0$ , the IFS final result ( $\lim_{n \rightarrow \infty} x_n = A$ ) it is always the same.  $A$  is called a Fractal set, being also known as the Attractor of  $f(x)$ .

Formally, an IFS is a finite set of contraction mappings on a complete metric space. Therefore, we can create iterative processes using simultaneously multiple contraction functions  $f_i(x)$ , so-called Hutchinson operators.

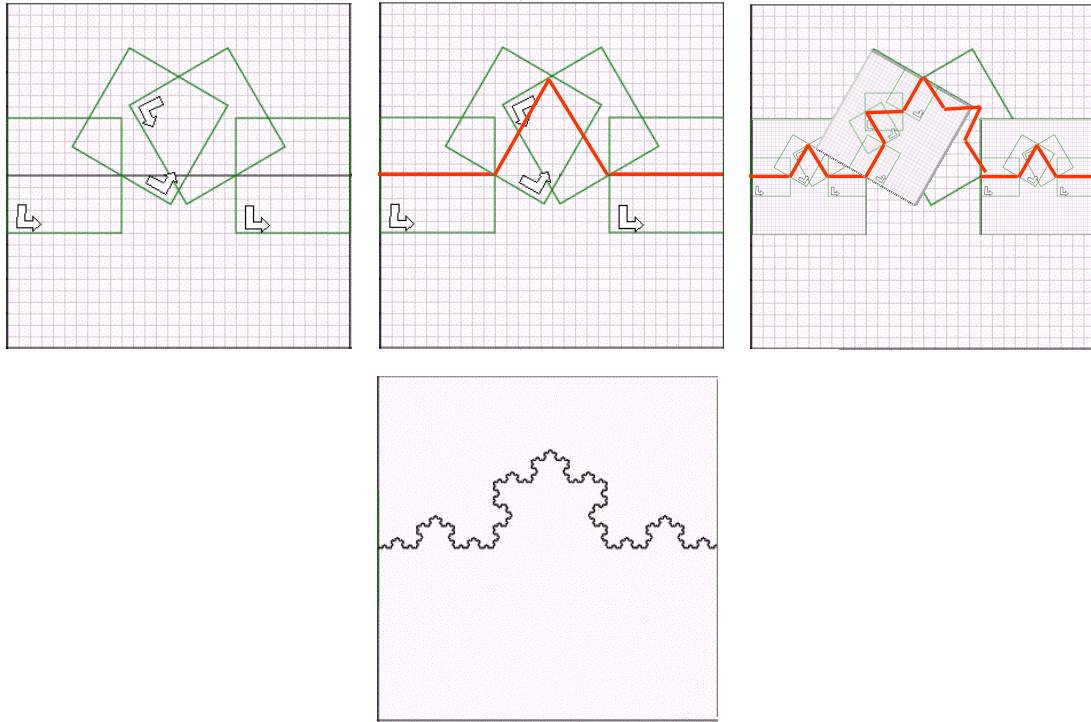
In Figure 4.2, the famous Von-Koch's fractals curve construction process is shown. Initially, the IFS input  $x_0$  is a line segment. The IFS is formed by 4 "scale-down+rotate+displace" functions (shown as green boxes), which transform the original segment in 4 new segments belonging to  $x_1$ . For each output segment, the same process is repeated. At the limit, we obtain a fractal curve of infinite length known as Von-Koch's curve.

<sup>3</sup> A mapping  $f$  is a contraction of the metric space  $X$ , provided that there is a constant,  $c$ ,  $0 \leq c < 1$ , such that for all  $x, y$  in  $X$ , and  $d(x)$  is the distance function, is accomplished that:

$$d(f(x), f(y)) \leq c \cdot d(x, y)$$

Where  $c$  is called the contraction factor



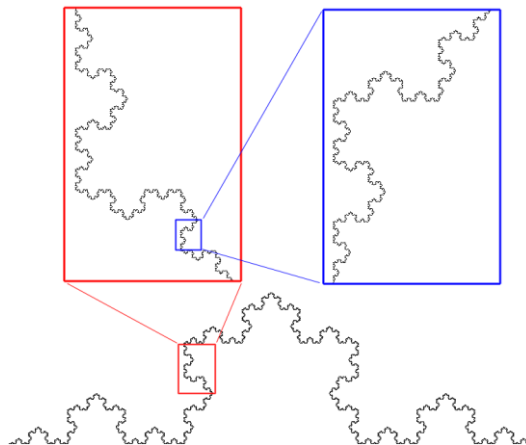


**Figure 4.2: IFS Construction of the fractal Von Koch's curve.**

Usually, for most renowned fractals,  $f(x)$  also accomplishes another property, the transformation is linear, and preserves the angles. Fractals created by nonlinear IFS are a specific area of research. We also can speak of statistical fractals in those cases where  $f(x)$  is not deterministic (i.e. could be modeled as a stochastic process), leading to another active branch of fractal geometry.

## 4.2 Classic Fractals and Self-Similarity

Self-similarity is one of the most important fractal properties, deeply related to the IFS process. The word self-similarity is quite self-explanatory; it defines those structures that look very similar when looked at different scales. This means that when a part of the structure is removed and compared with the whole, they are the same. If we use the Von Koch curve we have already built on previous section, we can easily notice self-similarity:



**Figure 4.3: Self-similarity of Von Koch curve**

The comparison scale ratio at which self-similarity occurs is an intrinsic property of each fractal set.

Fractal theory was build on top of the foundations of already-known mathematical sets, which defy the universal Euclidean rules. Those sets were characterized for its irregularity, in contrast of smooth regular curves and surfaces, and its self-similarity. Curiously, most of the classical fractal sets were known much time before fractal geometry was stated.

#### 4.2.1 Cantor set

The Cantor set was published in 1883, and has played an important role in many branches of mathematics, being the essential skeleton behind many other fractals. Building a Cantor Set is quite straightforward. We start with an initial set,  $X_0$ , formed by the unit interval  $[0, 1]$ , comprising the real numbers such that  $0 \leq x \leq 1$ . Cantor sets are obtained by iteratively applying a simple transformation rule:  $X_{n+1}$  set is obtained by removing the middle third of  $X_n$  intervals. Hence,  $X_1$  will be  $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ , and  $X_2$  subdivides the two previous subintervals, obtaining  $[0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$ . Cantor set is the limit of the sequence when  $n$  tends to infinity.



Figure 4.4: First iterations for Cantor set construction

Cantor Set has some properties that make it peculiar:

- It is an infinite and uncountable set. Its cardinality is the same as  $[0, 1]$
- As expected, it is self-similar. The set contained in  $[0, \frac{1}{3}]$  is an exact copy of the entire cantor set scaled down by  $1/3$  factor.

### 4.3 Deterministic and Natural fractals

The fractals sets which have been already presented are known as deterministic or perfect fractals. They are, in fact, mathematical devices far from real objects. In the other side, it has



Figure 4.5: Romanesco-Cauliflower displays a clear self-similarity

been observed that multiple nature phenomena exhibit in some way fractal properties, but not in a fully deterministic manner. It was at 1982 when Benoît Mandelbrot put in relation the fractal mathematics with plenty of natural structures in his reference book [5]. Classical examples of natural fractals are broccoli or cauliflower vegetables, trees and their branching structure, coastal-lines, lightning's paths or vessel of circulatory systems (arteries and veins). Other interesting examples related with biological growth may be found at [6]. A famous citation of

Mandelbrot may be enlightening:

*“Clouds are not spheres, mountains are not cones, and lightning does not travel in a straight line. The complexity of nature's shapes differs in kind, not merely degree, from that of the shapes of ordinary geometry, the geometry of fractal shapes”*

However, natural or statistical fractals show non-exact self-similarity. They are non exact, since the replica may be noisy or present some degree of distortion. Additionally, there is a scale range-limitation, caused by the inherent resolution limitation of the natural structures, and the maximum size limitation defined by the object itself. In other cases, the fractal behavior is limited to certain range of scales, modeling only partially the object.

Those effects are worsened by the fact that studying natural devices requires data sampling, which limits even more the maximum resolution available.

#### 4.3.1 Euclidean and Topological Dimension

---

The Latin word *dimensio* means *measure*. Dimension is an evasive concept but it could be generally defined as a characteristic number associated with a metric space. Mathematics has developed multiple definitions of dimension, being each of them applicable to certain group of problems and situations. One of the most familiar dimensions is the Euclidean,  $D_E$ , which takes into account the space occupied by an object. Euclidean geometry considers a structure one-dimensional if it is contained on a straight line, two-dimensional if it is embedded on a plane and three dimensional if it is a volume in space. Points are dimensionless, in other words, their dimension is 0.

Topological dimension  $D_T$  is defined regarding the way in which an object can be divided. The point is dimensionless, since it cannot be further split. A curve is one-dimensional since it could be divided using zero-dimensional shapes (points). Surfaces are two-dimensional since we need one-dimensional curves for dividing them. In some cases, topological and Euclidean dimensions coincide, is the case of a point ( $D_E=D_T=0$ ) or a straight line ( $D_E=D_T=1$ ). In others cases, it is different, i.e. a curve ( $D_E=2, D_T=1$ ) or a curved surface ( $D_E=3, D_T=2$ )

Returning to our running example, the Van Koch curve, it is easily demonstrated that has infinite length. Length is used to measure sets of dimension one, while area measures objects of dimension two. Van Koch curve has a dimension larger than one but smaller than two, hence a dimension should be allowed to be a fraction. In fact, “fractal” derives from the Latin word *frangere*, which means ‘to break’.

It is clear that classic dimensions are not suitable when dealing with fractal sets, next sections describe other dimension that will permit fractal set characterization. All of them are based in Mandelbrot’s fractal dimension, which in turn was motivated by the foundational Hausdorff dimension. Although the Hausdorff dimension is the most general result, it will not be covered here since its calculation is not feasible from a computational point of view.

#### 4.3.2 Self-similar dimension


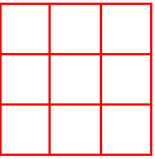
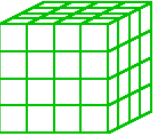
---

Self-similar dimension,  $D_s$ , is a concept applicable only to deterministic self-similar fractals, with a well known IFS and a regular transformation mapping.


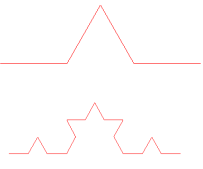
Given a certain IFS, and its Hutchinson operators, let be  $r$  the reduction factor of the IFS, and  $N$  the number of scaled-down pieces into which any IFS input structure is going to be divided in the current iteration step, i.e. the number of Hutchinson operators. Self-similarity dimension is defined as the power-law relationship between  $N$  and  $r$ :

$$N = r^{D_s} ; \quad \text{or equivalently} \quad D_s = \frac{\log(N)}{\log(r)}$$

When applied to non-fractal, but self-similar Euclidean objects, such as lines, planes or cubes,  $D_s$  obtain values that coincide with  $D_T$ . For fractal objects, as expected, a non-integer dimension is obtained.

	$r$	$N$	$N=r^D$	$D_s$
 line	5	5	$5^1$	1
 square	3	9	$3^2$	2
 cube	4	64	$4^3$	3

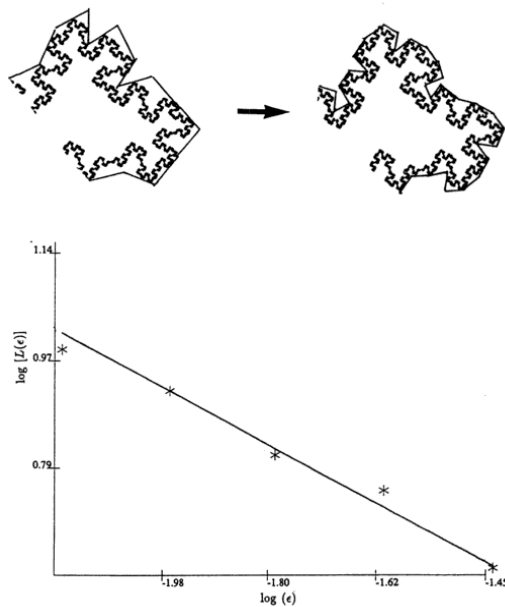
	$D_T$	$r$	$N$	$D_s = \frac{\log N}{\log r}$
 Cantor Set	0	3	2	0.63
	1	3	4	1.261

$D_s$  for Euclidean objects

$D_s$  for deterministic fractals

Figure 4.6: Similarity dimension calculation

### 4.3.3 Compass-Dimension



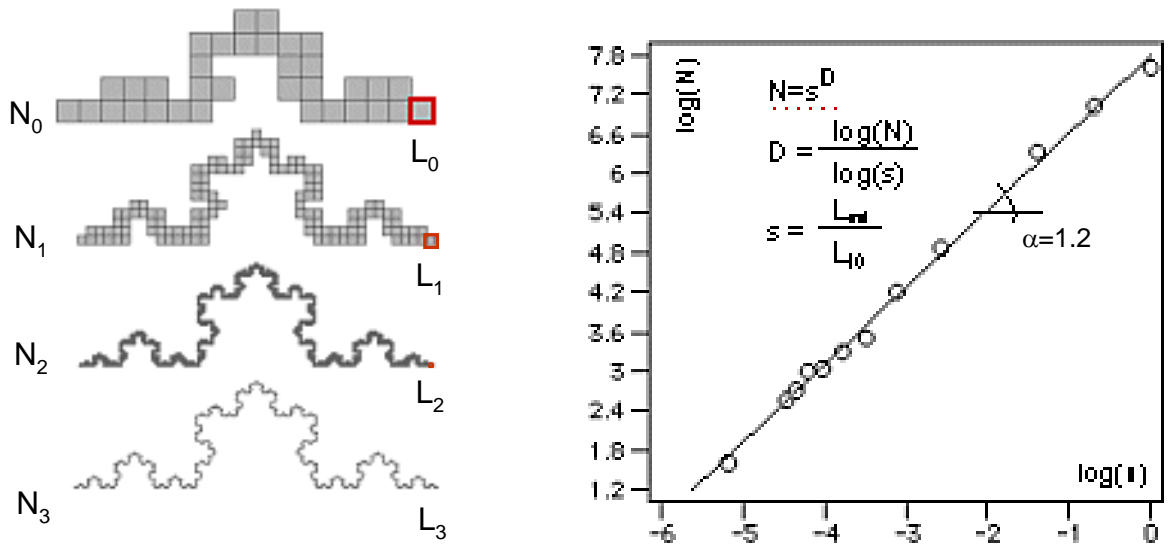
Compass dimension is a simple technique which permits obtaining a fractal dimension for deterministic and natural fractal curves. The idea is quite simple: with the help of ruler and compass, we will perform some length measurements of the curve. An initial compass distance  $s$  is selected and, starting from a point, the entire curve is walked with the compass, counting the number of required steps,  $n$ . Hence, the length of the curve at the scale  $s$  is  $l=n \cdot s$ . The process is repeated at different compass distances, reducing  $s$  each time. Since the fractal dimension is the exponent of the power-law involving scale and length, one of the typical tools for estimating such exponent is the log-log plot. In this case, we represent in a plot the  $\log(l)$  versus  $\log(1/s)$ . The exponent of the power law coincides with the slope of the line obtained as result of linear regression of the samples.

Figure 4.7 : Obtaining Compass dimension. Above: calculating length at different scales using multiple compass sizes. Below: representation of the size against the compass' arc size. Note that the dimension is the negative slope of the graph.

#### 4.3.3.1 Box-Counting dimension

When the analyzed set is not a deterministic fractal and its structure is not suitable to be traced with ruler and compass, a new approach to the problem may be taken. The box-counting algorithm studies the dimension by matching a grid against the set. At each step, the number of grid boxes containing set elements is counted. Afterwards, the grid resolution is increased (typically doubled, though other scale factors are possible), and the counting process is repeated. The process is iterated while the data set resolution allows it. The box-counting dimension is obtained as the exponent of the power-law relationship of the grid's resolution ( $s$ ) versus the box count ( $N$ ) at a given grid's resolution. Again, a log-log-plot is utilized, the graph of the  $\log(N)$  vs  $\log(s)$  is represented, and approximated by linear regression. The line slope coincides with the Box-counting dimension  $D_B$ .

$$(4.1) \quad D_B = \lim_{B \rightarrow 0} \frac{\log(N)}{\log(1/B_{size})} = \lim_{s \rightarrow \infty} \frac{\log(N)}{\log(s)}$$



**Figure 4.8: Box-counting dimension on Von Koch curve. Left: increasing resolution covering grids. Right: log-log plot and slope calculation.**

Box counting and its variations is the most used technique dimension in fractal measurements, since it has an easy computer implementation, and it is easy scalable to more-than-2 dimensional spaces. In the other hand, obtaining reliable results require that the finest grid resolution must be sufficient to guarantee that the log-log-plot's slope tends to the fractal dimension.

#### 4.3.3.2 Differential box counting

Until this moment, the methods to determine fractal dimension (FD) already discussed are applicable to binary sets. When dealing with grayscale images, it is required to use new approaches. The first and simplest option consists of representing the image as a 3D surface. The process is quite simple, given an  $M \times M$  sized image, the algorithm works by mapping  $(x,y)$  pixel coordinates to  $(x,y)$  vertex's surface coordinates, and the intensity of the image pixel is mapped to the  $z$  vertex coordinate. At this moment, a 3D version of the well-known box-counting is applied. The 3D space is partitioned in cubes of size  $r \times r \times r$ . Regarding the  $z$  coordinate, in order to apply a homogeneous scaling, it is required to measure the image intensity using a unit derived from the quotient of the maximal intensity value  $I_{max}$ , and the image width  $M$ . I.e., an alternative image intensity  $I'$  will be utilized:  $z = I'(i,j) = I(i,j) * M / I_{max}$ . In order to count the number of boxes required to cover the whole surface, the differential

approach is used, which relies on utilizing the maximum and minimum intensity value within the cell, and applying the formula (4.2):

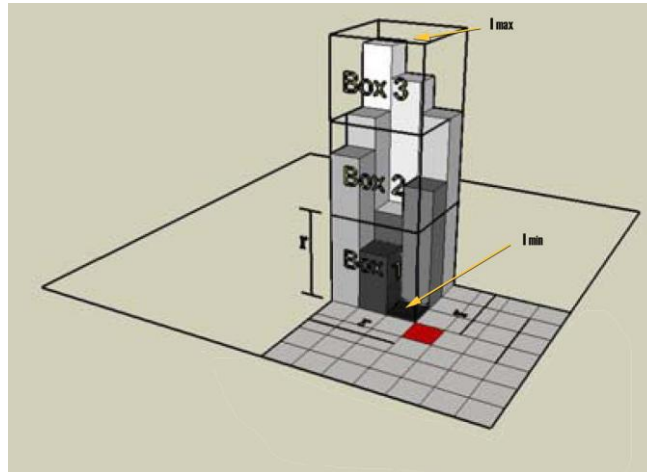
$$(4.2) \quad N_{box} = \left\lfloor \frac{I_{max} - I_{min}}{r} \right\rfloor + 1$$

$$(4.3) \quad N_r = \sum_{i,j} \left( \left\lfloor \frac{I'_{max}(i,j) - I'_{min}(i,j)}{r} \right\rfloor + 1 \right)$$

That system of box counting is called **differential box counting** or DBC, given the differential nature of the calculus.

Finally, the total number of boxes per resolution is counted as (4.3), and fractal dimension is obtained by linear regression of  $\log(N_r)$  versus  $\log(r)$ .

As presented in [14], it is important to subtract and divide before flooring the result, and avoid calculating  $\left\lfloor \frac{I_{max}}{r} \right\rfloor - \left\lfloor \frac{I_{min}}{r} \right\rfloor$ , since we will suffer from quantization errors.



**Figure 4.9: Differential Box Counting.** Box of size  $r$ , and its  $z$  column. Selected box has a count of 3 boxes.

## 4.4 Multifractal theoretical overview

### 4.4.1 Introduction

The approach of Fractals is based in the study of sets. Elements may belong or not to a set, following a binary true or false relationship. However, this approximation is not adequate for representing most of the natural phenomena, where studied parameters may have larger dominions, ranging from a limited set of values to unlimited continuous levels. That is the motivation for the appearance of the *measure* concept. The idea behind multifractals is to extend fractal self-similarity from sets to measures.

Another keystone of multifractal theory is the inexactitude of self-similarity in natural processes. The similarity observed at several scales is far to be exact in most of the natural phenomena. Additionally, in the same process we may notice similarity at different scales, located in different areas. This means that multiple fractal sets lie interwoven, each one with their own scaling behavior. That is the origin of multifractal word.

The theoretical contents found in this section are covered in the references [7][8][22]

### 4.4.2 Definitions

#### 4.4.2.1 Measure

A measure of a bounded subset  $S$  of  $\mathbb{R}^n$ , expressed as  $\mu(S)$ , can be often considered as a sort of mass distribution. A metaphor may help understanding the concept; think of spreading a little handful of sand in a bounded surface, the mass sand distribution on the surface in a given point corresponds to the measure.

#### 4.4.2.2 Local behaviour: Holder Exponent

First phase of multifractal analysis is to study the local behaviour of our data. To do so, we will be using a “magnifying lens” focused in a certain box  $B$  of side  $\varepsilon$ , and centered on an element of the set  $x$ . Given that settings, the coarse Hölder exponent [22] is defined as the exponent that regulates the power law of measure and scale. If limit when  $\varepsilon$  tends to zero is taken, we speak of fine Hölder exponent, which is a sort of local dimension.

$$(4.4) \quad \alpha_c = \frac{\log(\mu(B(x, \varepsilon)))}{\log(\varepsilon)}$$

$$(4.5) \quad \alpha_f = \lim_{\varepsilon \rightarrow 0} \frac{\log(\mu(B(x, \varepsilon)))}{\log(\varepsilon)}$$

The Hölder exponent is calculated for the whole data set. Depending on the nature of the studied process, the range of values  $[\alpha_{\min}, \alpha_{\max}]$  may vary. In some cases it is restricted to a small value interval, in others it can range from 0 to infinity.

#### 4.4.2.3 Global behaviour: Multifractal spectrum

Once Hölder exponents have been calculated, we place an  $\varepsilon$ -resolution grid onto the analyzed set. Next, for each  $\alpha$  value, we count the number of boxes of size  $\varepsilon$  having coarse Hölder exponent equal to that  $\alpha$ , obtaining  $N_\varepsilon(\alpha)$ . The coarse multifractal spectrum [22] is defined as:

$$(4.6) \quad f_c(\alpha) = \lim_{\varepsilon \rightarrow 0} \frac{\ln^+(N_\varepsilon(\alpha))}{\ln(\varepsilon)}.$$

---

<sup>4</sup>  $\ln^+(x) = \max(\ln(x), 0)$

When analyzing a data set,  $f_c(\alpha)$  gives a fractal dimension for each one of the subsets formed by those elements sharing a certain singularity pattern determined by  $\alpha$ .

Obtaining the fine multifractal spectrum may require calculating the Hausdorff dimension, which is computationally unaffordable, and it is out of the scope of this work.

#### 4.4.3 Methods for multifractal spectrum calculation.

---

##### 4.4.3.1 Method of the Histogram

The direct approach for multifractal spectrum consists of four steps [7]:

- 1) For each set element, calculate the coarse Hölder exponent at maximum resolution available, using the expression (4.4).
- 2) Discretize the  $\alpha$  range in  $\Delta\alpha$  sized segments, giving a set  $\{\alpha_1, \dots, \alpha_i\}$  of discrete values
- 3) Choose a set of box-grids with increasing resolution.
  - a) For each grid find of scale  $\epsilon_i$ ,
    - i) For each  $\alpha_i$ 
      - (1) Calculate  $N_\epsilon(\alpha, \Delta\alpha)$ , i.e. we count the number of boxes having Hölder exponent between  $\alpha_i$  and  $\alpha_i + \Delta\alpha$
- 4)  $f_c(\alpha)$  is derived from equation (4.6) using linear regression: for each  $\alpha_i$ , the slope of  $\log(N_\epsilon(\alpha_i, \Delta\alpha))$  against  $\log(\epsilon_i)$  equals  $f_c(\alpha_i)$

##### 4.4.3.2 Method of moments

Direct method is time consuming from a computational point of view. A good alternative consist of using the method of moments, which is based in the concept of generalized dimension[7].

Generalized dimension  $D_q$  corresponds to scaling exponents for the  $q$ -th moments of the measure. It provides an alternative description of the singularity measure. Once again, covering the support of the measure with boxes of size  $l$ , and defining  $P_i(l)$  to be the probability of the  $i$ -th box, a series of exponents parameterized by  $q$  according to:

$$(4.7) \quad \sum_i [P_i(l)]^q \approx l^{(q-1) \cdot D_q}.$$

The exponent is calculated from the following expression:

$$(4.8) \quad D_q = \frac{1}{(q-1)} \lim_{l \rightarrow 0} \frac{\log \sum_i [P_i(l)]^q}{\log(l)}.$$

Certain values of  $D_q$  have special significance;  $D_0$  coincides with the box-counting dimension.  $D_1$  is related with the information theory developed by Claude Shannon, being known as Information dimension.  $D_2$  is called correlation dimension.

Generalized dimension could also be described using the parameter  $\tau(q)$  alternatively to  $D_q$ . Both concepts are easily obtained from each other:  $\tau(q) = D_q \cdot (q-1)$ .



There is a strong relationship between the multifractal spectrums and the generalized dimension.  $\tau(q)$  is related with  $f(q)$  through Legendre transform, which is described by the following equations:

$$(4.9) \quad \frac{\partial}{\partial q} \tau(q) = \alpha(q).$$

$$(4.10) \quad f(\alpha(q)) = q \cdot \alpha(q) - \tau(q).$$

Therefore, one of the usual methods for obtaining the multifractal spectrum starts by obtaining the generalized dimension and, later on, utilizing the described equations to obtain  $\alpha$  and  $f(\alpha)$ .

#### 4.4.3.3 Direct determination of the $f(\alpha)$ singularity spectrum

Chaabra and Jensen described a method [9] for calculating the multifractal spectrum without requiring the usage of Legendre transform. The following procedure is applied:

- a) We cover the experimental data set with boxes of size  $l$ , and compute the probability of each of the boxes  $P_i(l)$ . Afterwards, normalized measures for multiple moments  $\mu_i(q, l)$  are obtained using:

$$(4.11) \quad \mu_i(q, l) = \frac{[P_i(l)]^q}{\sum_j [P_j(l)]^q}.$$

$q$  acts as an amplifier lens that permits exploring different regions of the measure. For  $q > 1$ , those singular measures are amplified, for  $q < 1$ , most frequent measures are analyzed.  $q = 1$  replicates the original measures.

- b) Finally,  $f(q)$  and  $\alpha(q)$  are obtained from:

$$(4.12) \quad \alpha(q) = - \lim_{l \rightarrow 0} \frac{\sum_i \mu_i(q, l) \cdot \ln(P_i(l))}{\ln(l)}$$

$$(4.13) \quad f(q) = - \lim_{l \rightarrow 0} \frac{\sum_i \mu_i(q, l) \cdot \ln(\mu_i(q, l))}{\ln(l)}$$

In both cases, it is required to perform a linear regression for each  $q$  and scale combination, in order to obtain the approximation of the limits.

#### 4.4.3.4 Multifractal from Discrete Wavelet Transform (DWT)

Last approach to multifractal spectrum calculation is done through the usage of the Discrete Wavelet Transform. Intuitively, it seems quite natural to find some similarities between both mathematical devices. DWT allows the exploration of a signal at multiple resolutions and locations, which is deeply related to multifractal analysis.

By using sliding wavelets instead of boxes, we get rid of the grid position dependence that might mask singularities or perturb the estimation of Hölder exponents. The statistical approach for obtaining multifractal spectrum using wavelets is called Wavelet Transform Modulus Maxima (WTMM). As shown in [10], the mathematical background is complex, but method may be summarized in a few non-trivial steps:

- 1) Calculate the DWT coefficients  $T$ , and obtain its modulus  $M$ .

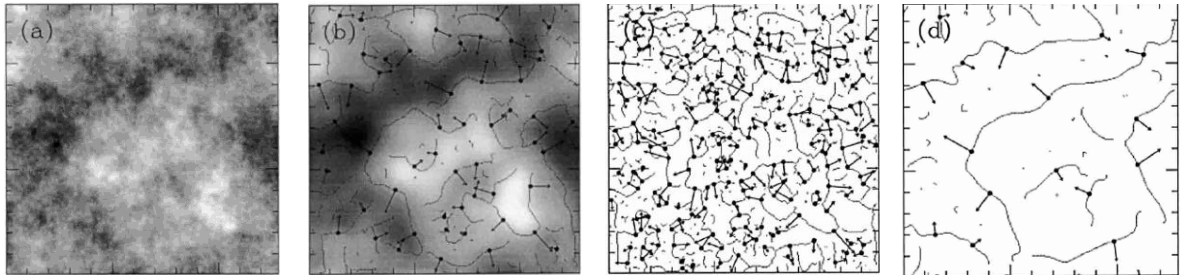
$$(4.14) \quad \mathcal{M}_\psi[f](\mathbf{b}, a) = |\mathbf{T}_\psi[f](\mathbf{b}, a)|$$

- 2) Local maxima chains of  $M$ , so-called WTM skeletons, are calculated at different scale resolutions  $a$ . Maxima from different resolutions levels are linked using proximity criteria.
- 3) Utilize wavelet transform skeleton to compute the partition function  $Z(q, a)$

$$(4.15) \quad \mathcal{Z}(q, a) = \sum_{\mathcal{L} \in \mathcal{L}(a)} \left( \sup_{(\mathbf{x}, a') \in \mathcal{L}, a' \leq a} \mathcal{M}_\psi[f](\mathbf{x}, a') \right)^q$$

Being  $\mathcal{L}(a)$  the set of all maxima lines that exist at the scale  $a$  and which contain maxima at any scale  $a' < a$ .

- 4)  $\tau(q)$  is obtained from the slope of linear regression of the plot  $\log(Z(q, a))$  vs  $\log(a)$
- 5) Finally, obtain  $f(\alpha(q))$  utilizing Legendre transform.



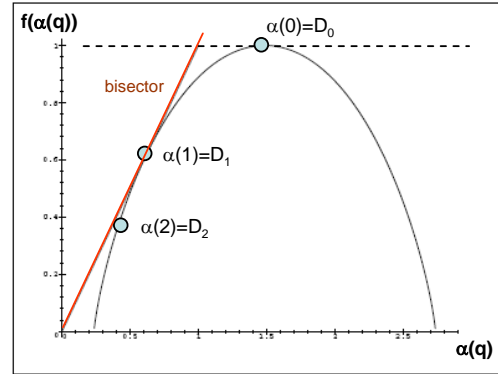
**Figure 4.10:** (a) 32 grey-scale image. In (b) grey scale coded WTM. From (b) to (d), lines are maxima skeletons of WTM at varying resolution  $a$ . Dots show local maxima and its direction. Resolutions for (b), (c) and (d) are,  $a = 2\sigma_w$ ,  $a = 2^{0.1}\sigma_w$ , and  $a = 2^{1.9}\sigma_w$ , respectively.

It could be noticed that  $Z$  takes into account the local maxima of the maximal lines at a certain scale  $a$ . Those points are, in fact, singularity points which could be thought as areas with the same holder exponent.

#### 4.4.4 Basic properties of the multifractal spectrum

Some of the fundamental multifractal spectrum properties need to be reviewed:

- 1) The spectrum has an interesting relationship with the internal bisector of the axis  $\alpha$  and  $f(\alpha)$ :
  - a) The spectrum is below the bisector of the axis:  $f(\alpha) \leq \alpha$
  - b) The spectrum touches the bisector line at  $q=1$ .  $\alpha(1) = f(\alpha(1))$
- 2) The spectrum reaches its maximum  $D_0$ , at  $q=0$ :  $\alpha(0)$ .  $f(\alpha(q)) \leq f(\alpha(0))$
- 3) For most general multifractal processes generated by multiplicative cascades, spectrum curve resembles a concave bell below the bisector.



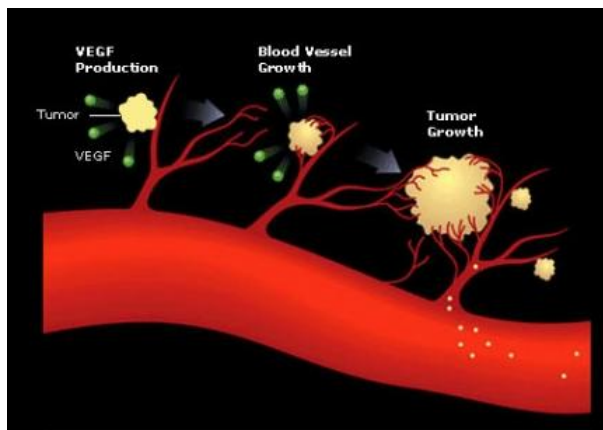
**Figure 4.11: Typical multifractal spectrum, below the bisector and tangent at  $\alpha(1)$**

## 5 PREVIOUS WORK.

### 5.1 Fractals and Cancer. Is there a relationship?

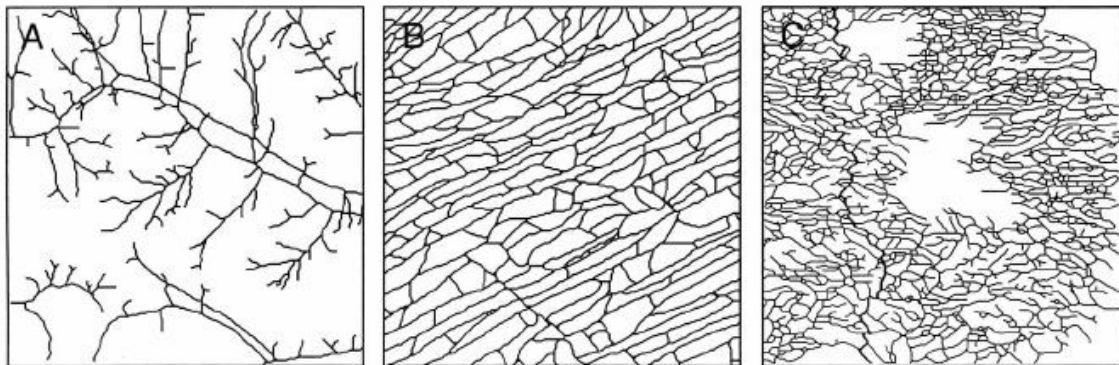
#### 5.1.1 Angiogenesis and statistical growth processes

Tumor vasculature is known to be more chaotic in appearance than normal vasculature. This is related to one of the most critical events in tumor progression: the angiogenesis. Tumoral cells emit signalling molecules to the host tissues, which stimulates the generation of proteins required for generating new blood vessels. Those vessels will be responsible for delivering nutrients and oxygen, and removing detritus of tumor cells.



**Figure 5.1: Blood vessel generation (angiogenesis) due to the signalers emitted by the tumor.**

Experimental settings with mice have demonstrated that tumor have quite different vascular densities and growth characteristics, furthermore, the fractal dimensions of tumors vessels [15] has been found to be  $1.89 \pm 0.04$  whereas normal arteries and veins yield dimensions of  $1.70 \pm 0.03$



**Figure 5.2: Skeletonized images of vascular networks. A) normal vessels. B) normal subcutaneous capillars. C) Tumoral vasculature**

Tumor growth has been deeply studied, and its pathological morphology has been described using Euclidean parameters such as size, diameter and volumes, but the tumor structure is difficult to characterize in that manner. Fractal geometry opens a new observation window, which permits quantifying some of those morphological characteristics which physicians rely on to describe malignancies, such as the irregular border or unstructured vascular growth.

In fact, it has been found that using some mathematical models known as *statistical growth processes*, tumor and vascular growth can be reproduced. Furthermore, such models are also

known to produce fractal structures. Following table describes three growth algorithms which are relevant in cancer study.

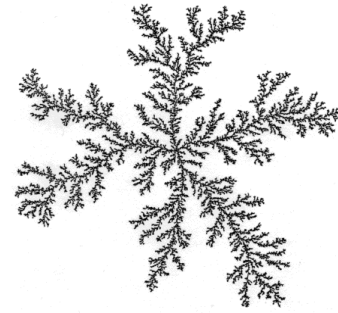
#### 5.1.1.1 Eden Growth

EG is a statistical growth process that works on a planar grid of cells. Initially, a single cell is infected. Next site to be infected is chosen at random from among the nearest neighbors of the first site. Growth continues by randomly infecting the neighbors of growing cluster as it expands.

#### 5.1.1.2 Diffusion Limited Aggregation (DLA)

DLA, is a statistical growth model based on the Eden growth. EG model is enriched with the addition of a scalar field which represents a substance that is diffusing toward the cluster. That substance may simulate the presence of a nutrient, or a tumoral signaler.

In DLA the probability of growth of the candidate cells is not spatially uniform. The cells where the gradient of the diffusing substance is larger have higher probabilities to be selected. This process generates dendritic-like structures that have been suggested as model of growth of healthy arteries. The fractal dimension of a DLA cluster as a whole is known to be about 1.71 in a plane with a minimum path dimension of 1.00. The minimum path dimension is the fractal dimension of the shortest path crossing the fractal structure.



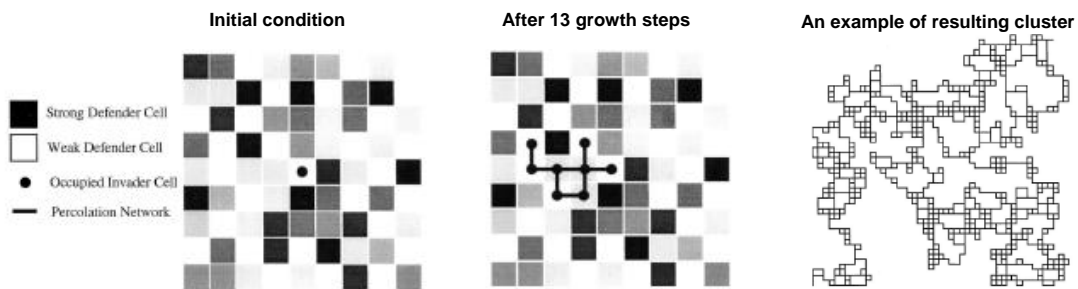
**Figure 5.3: DLA generated figure**

#### 5.1.1.3 Invasion Percolation

In several science fields, percolation is related to the movement and filtering of fluids through porous materials. This process may be simulated using a simple algorithm:

- a) A grid of cells is created, and a strength value is assigned randomly to each cell.
- b) A cell is selected as seed for a growing invader cluster.
- c) At each time step, the current invader structure is enlarged with the addition of the neighboring cell with the lower strength.

Resulting percolation clusters showing up holes of many sizes, leading to a fractal dimension of 1.89 in a plane for the cluster and a minimum path dimension of 1.13. It is important to remark that that the observed fractal dimension of the tumor vasculature corresponds to those produced by an invasion percolation. Invasion percolation results largely depend on the randomness of the background expansion network, which in the case of tumors is supplied by the host tissue.



**Figure 5.4: Invasion Percolation algorithm**

### 5.1.2 Imaging using contrast agents

---

Because of the peculiar vascular architecture of tumor vessels, the movement of blood-borne substances such as imaging tracers is different from normal veins or arteries. When an imaging tracer is applied, the substance does not reach the whole tumor volume as expected. This is caused by the tortuosity of the vessels. Using MRI imaging it has been shown that perfusion of the imaging marker in the tumor behaves like in a percolation network, flowing rapidly in some places, flowing slowly in others and not flowing at all elsewhere. It also has been found that the elimination of a tracer from a tumor has a prolonged tail provoked by the heterogeneous transportation of vessels.

### 5.1.3 Fractals and Cancer

---

Regarding the applicability, a lot of studies established a clear relationship amongst cancer and fractals. Remarkable efforts have been done in the field of mammograms, where the fractal analysis tools have demonstrated to be really useful.

In the other hand, it could not be generalized that tumors introduce more irregularity and imply the rising of fractal dimension. Counter-examples have been found for the rule. In the case of metastatic bone marrow<sup>5</sup> lesions, where the affected tissue shows a loss of fractal structure. Simulations performed with automata models showed that the fractal dimensions on the tumor border depend on a complex interplay of parameters such as tumor motility, the supportive framework of the organ or structure, the hormone signalers proliferation due to intra-cell communication or endocrine cells transmissions, tumoral destruction ratio, etc. Nevertheless, fractal methods have shown to be a powerful tool for examining the mechanistic origins of the pathology. Literature covers several examples of its usage for different tumoral pathologies such as pulmonary, prostatic or breast cancer.

## 5.2 FA: Fractal Analysis on medical images

This section covers some Fractal analysis techniques and its application to tumor detection and classification on medical images. The existing bibliography is abundant but rather heterogeneous, since this field has been covered by a multidisciplinary range of specialists such as biologists, medical practitioners, physics experts and all sort of engineers.

### 5.2.1 Study cases

---

In [11], fractal analysis is applied to characterization of prostate cancer in MR images. For such purpose, T2-weighted images (T2WI) of 55 patients were used. T1-weighted images are not suitable for prostate cancer detection, since they could not depict prostate zonal anatomy. From 55 cases, 28 were biopsy-proven cancers, and 27 were normal.

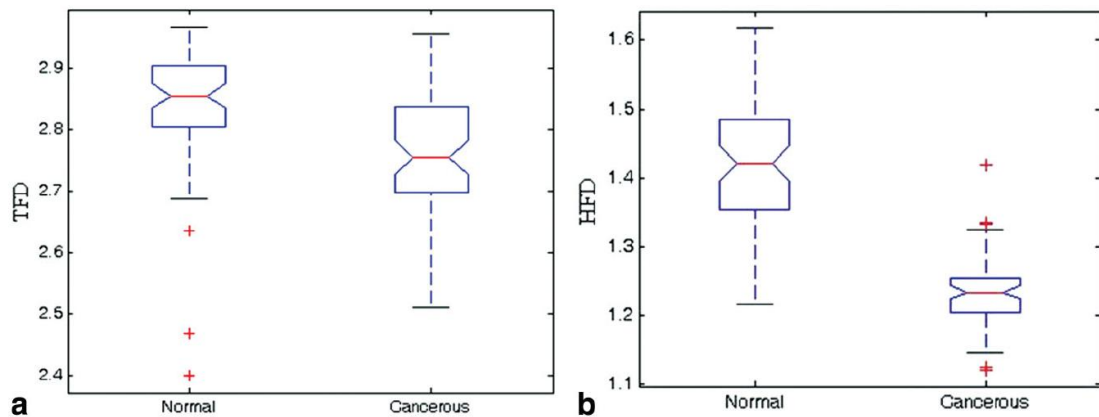
The study demonstrates a statistically significant relationship between texture fractal dimension (TFD) and histogram fractal dimension (HFD) with benignity or malignancy of explored tissues. The experiment worked on a collection of manually selected 12x12 images of prostate regions of interest (ROI) selected from the group of MR images. The image size was small enough to guarantee that the tissue selected was fully inside the affected regions for those positive cases.

The set of images were classified using TFD and HFD techniques. TFD is another name for differential box counting, already introduced, while HDG deals with the study of the fractal dimension of the image histogram. The histogram is a classical image processing analysis tool which shows the intensity distribution curve. Its fractal dimension is calculated using a 2D DBC.

---

<sup>5</sup> The substance inside bones which produces blood cells

Next figure summarizes the results obtained when trying to differentiate normal and cancerous tissues. It could be seen that, in spite of both techniques are able to classify tissues, HFD performs fairly better.

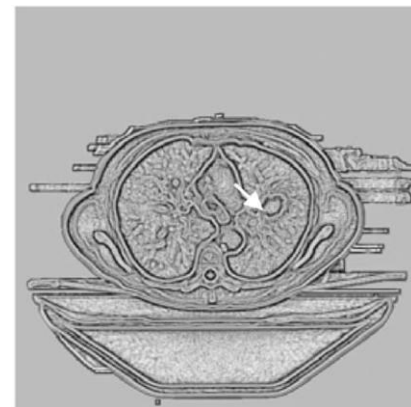


**Figure 5.5: Box-plots of statistical distributions for TFDs (a) and HFDs (b), in the normal ROIs and cancerous ROIs. Red lines in the boxes represent median values, blue boxes cover the IRQ (50% of the samples). Whiskers extending from the boxes show the extent of 1.5 x IQR of the data (99.3% of the samples).**

ROC analysis confirms that HFD achieves better classification results than TFD, being their respective AUC<sup>6</sup> 0.966 versus 0.691. It is speculated that the reason for such a difference is that TFD is more sensible to spatial variation of intensity pixels. Since prostate cancer is multifocal in most of the cases, its spatial distribution sometimes could be similar to normal tissue. HFD is not sensible to spatial distributions, which makes it more accurate for that type of tumors.

.....

Another interesting utilization of fractal analysis is performed in [12], where its application to aggressive and nonaggressive lung tumors was proposed. In this case, the goal of FD analysis is not focused on not tumor detection, but in tumor staging. Even for experienced radiologists, it is difficult to diagnose in a CT whether a small-cells lung tumor is aggressive (stages 3,4) or not (stages 1,2), which is capital when deciding the most adequate treatment to the patient. The authors found that there is a high correlation between the fractal dimension and the stage of tumors. The higher the FD, the higher the stage.



**Figure 5.6: Fractal dimension Image. Arrow shows an early non-aggressive tumor**

The experimental data set was obtained using a contrast-enhanced computer tomography (CE-CT), which generates a group time-sequence images where contrast diffusion is depicted. The authors divided the work in three phases: transformation, selection of affected regions, and fractal analysis.

The transformation phase requires calculating a value of fractal dimension for each image pixel, generating a new image in the fractal dominium, the FD-image. For such a purpose, a scalable surrounding box around every input pixel is used, and its fractal dimension is calculated using DBC.

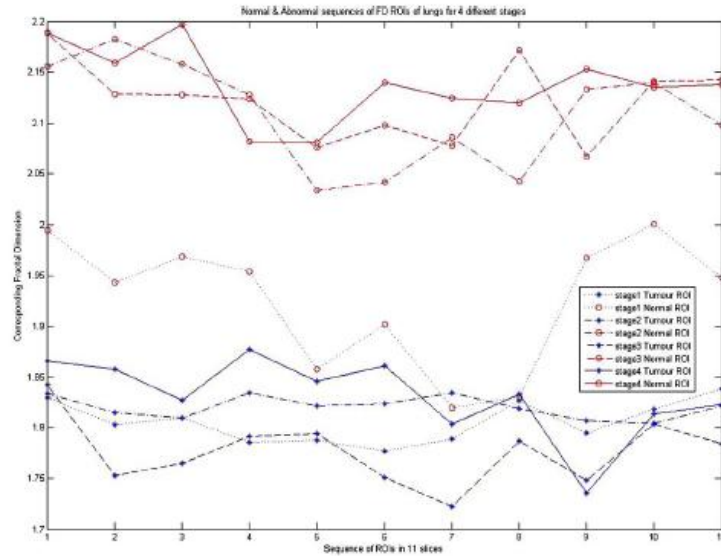
<sup>6</sup> AUC: Area under the ROC curve

The calculation of FD-images is a proper preprocess for CT images, easing human interactive selection of tumoral tissues. FD-images show a special sensitivity to irregular areas, and have the benefits of border enhancement. On these images, the physicians perform the selection phase by manually segmenting ROI images on lung's affected areas. Each selected area is displayed through the temporal series, allowing the operator to modify the position of the ROI in some temporal slices to correct the patient's possible respiratory motion.

Once the ROI images are selected, they are studied during the time series, calculating the averaged FD ( $FD_{avg}$ ) of each temporal snapshot. Each ROI is represented with an unique parameter: its time-maximum  $FD_{avg}$ .

The test has been run on 15 cases clinically rated, and thresholds for separating the stages 4 and 5 obtained, with a success rate of 83.55% for studied cases. FD showed strong correlation with corresponding standardized tumor uptake value obtained from FD-PET imaging.

Furthermore, the CT acquisition parameters, voltage and intensity, has been manipulated in order to study its effects on the FD calculation. It has been found that the standard deviation of FD for consecutive time-slices in a study area varies depending on the parameter set. That experiment has determined the combination of voltage and intensity that ensures the most precise measure.



**Figure 5.7: ROI's FD time evolution of four stage tumors in different stages and four normal cases. It can be seen that the upper area groups lung tumor sequences, and that average FD is proportional to the stage.**



## 5.3 MFA: Multifractals Analysis

### 5.3.1 Multifractal Analysis: an interdisciplinary tool.

Multifractal analysis was first introduced in physics, while studying the turbulence phenomenon. Since then, the technique has evolved, becoming a mathematical tool which has been applied in multiple scientific disciplines. Apart from biological structures, other multiple nature or human-driven phenomena show multifractal behavior, such as earth quake distribution, internet data traffic, stock market index evolution, etc. Two references are given below in order to get a glimpse of the wide spectrum of applications.

Surprisingly, it has been stated that WAN traffic presents self-similarity when observed at different time scales. Feldman et al. [16] demonstrated that due the inherent nature of some WAN traffic related protocols (TCP/IP), there exists a mathematical multiplicative process which can model consistently the traffics. A direct consequence of that result is that the traffic must show multifractal properties due the properties of multiplicative processes. Such models may allow generating more realistic synthetic traffic, inferring instant (low-scale time) traffic from statistical samples (high-scale time) and providing a better understanding of network traffic. For the analysis of the real traffic samples, Discrete Wavelet Transform method has been applied in order to obtain the multifractal spectrum of the one-dimensional traffic signals.

Another example of multifractal analysis is the characterization of soil particle-size distributions (PSD) [17]. PSD defines a combination of clay, silt and sand contents, which facilitates to group soil into classes. Multifractal analyses revealed PSDs with single (monofractal) and multifractal scaling. It was found that multifractal parameters were related to clay content. A clay content of 10% marked the threshold from single to multifractal scaling. For the experiments, multifractal spectrum,  $f(q)$  vs.  $\alpha(q)$ , was calculated as Chhabra and Jensen.

### 5.3.2 Multifractals applied to cancer detection and segmentation.

Huge efforts have been done in medical imaging in order to assist physicians in their fight with cancer. This chapter will focus in the applications of multifractals in breast cancers, which is one of the tumor classes with a higher impact and whose early diagnosis becomes determinant for treatment. For such purposes, multiple techniques have been developed in order to assist the radiologist and medicine practitioners in the diagnosis and treatment planning tasks. Diagnosis may be doubtful even for skilled radiologist, and support imaging tools can help on early detection, avoiding unnecessary biopsies (false positives) and help in a precise localization of affected tissues. Multifractal formalisms has demonstrated to be a powerful tool for both classification and segmentation tasks. Next section exemplifies each operation by studying two selected state-of-the-art papers.

#### 5.3.2.1 2D segmentation algorithm

One of the significant signs of breast cancerous activity is the presence of small mineral deposits in the tissue, also known as microcalcifications. Mammographies are the most extended imaging technique for detection and diagnosis, and allow detecting detect those microcalcifications in a non-invasive manner. However, some microcalcifications have a small size and are situated in low-contrast areas, which make the detection error prone.

Since multifractal analysis performs a study both at local and global scale, it is appropriate to detect the microcalcifications. For such purposes in [20] the multifractal analysis is adapted to the specificities of the case, in order to be able to detect and segment microcalcifications.

The approach relies on the calculation of multifractal spectra utilizing the histogram method, but adding some interesting tune-ups in order to make it more sensible to the target. First adaptation is to utilize a *capacity measure* instead of a simple measure. The simplest measure  $\mu(i,j)$  for a pixel is its grey level  $g(i,j)$ , which can be normalized dividing by the sum of grey levels of the

whole image. A capacity measure consist of the usage a kernel in a surrounding area  $\Omega$  around the pixel, which will give us a measure value, probably different than the pixel's own value. Multiple capacity measures are possible, but the authors have concentrated their efforts in the maximum and minimum capacity measures:

$$(5.1) \quad \text{Maximum capacity measure:} \quad \mu(i, j) = \max_{(i,j) \in \Omega} g(i, j)$$

$$(5.2) \quad \text{Minimum capacity measure:} \quad \mu(i, j) = \min_{(i,j) \in \Omega} g(i, j)$$

In fact, those measures resemble the morphological operators of dilation and erosion. Hence, the shape of  $\Omega$  determines the spatial behavior of the results. For microcalcification detection, the diamond shapes perform better than square or circular areas, and its size domain is 1, 3 and 5.

0.76	0.75	0.74	0.75	0.77
0.77	0.76	0.75	0.77	0.79
0.81	0.81	0.79	0.81	0.87
0.83	0.82	0.86	0.91	0.95
0.86	0.86	0.93	0.97	0.98

0.76	0.75	0.74	0.75	0.77
0.77	0.76	0.75	0.77	0.79
0.81	0.81	0.79	0.81	0.87
0.83	0.82	0.86	0.91	0.95
0.86	0.86	0.93	0.97	0.98

0.76	0.75	0.74	0.75	0.77
0.77	0.76	0.75	0.77	0.79
0.81	0.81	0.79	0.81	0.87
0.83	0.82	0.86	0.91	0.95
0.86	0.86	0.93	0.97	0.98

**Figure 5.8: Calculation of the maximum capacity measure for the pixel in the center of the image, using Diamond-shaped kernel and sizes 1,3 and 5. Dark grey is the measure value selected.**

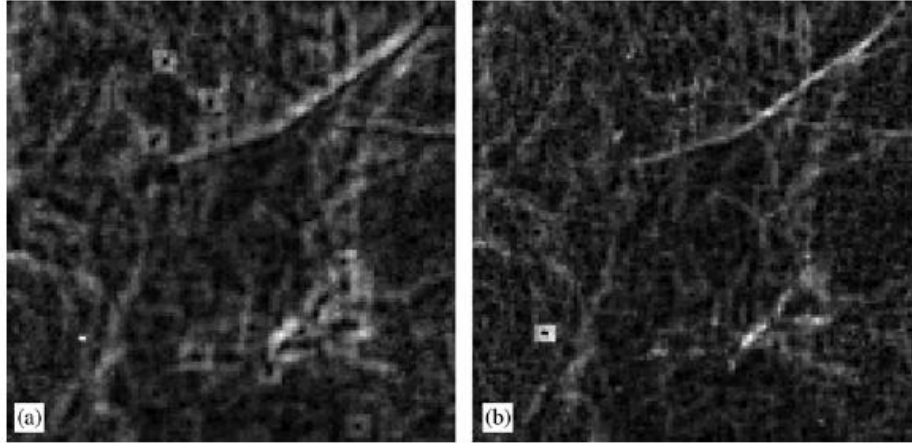
The multifractal is calculated in three phases:

1. The capacity measures are calculated
2. For each pixel, we consider a set of boxes of increasing size  $j$  and centered on the pixel. Domain of  $j$  is  $\{1,2,4,6,8,10,12,14,16\}$ . For each box, the coarse Hölder exponent ( $\alpha_c$ ) is calculated using (4.4). Following step is to approximate the fine Hölder exponent  $\alpha$  by the slope of the linear regression of  $\log(\alpha_c)$  versus  $\log(j)$ .

This step leads to an interesting result, the  **$\alpha$ -image**, which is an image whose pixel grey values correspond to  $\alpha$ , of course, adapting  $\alpha$  dynamic range to gray scale available range.

3.  $\alpha$  values are quantized and box counting is applied to calculate corresponding  $f(\alpha)$ , as explained in previous chapters. In a similar way, a  **$f(\alpha)$ -image** is obtained.

Several test show that maximum measure works better than the minimum, as could be appreciated in the Figure 5.9, where the  $\alpha$ -images are shown.

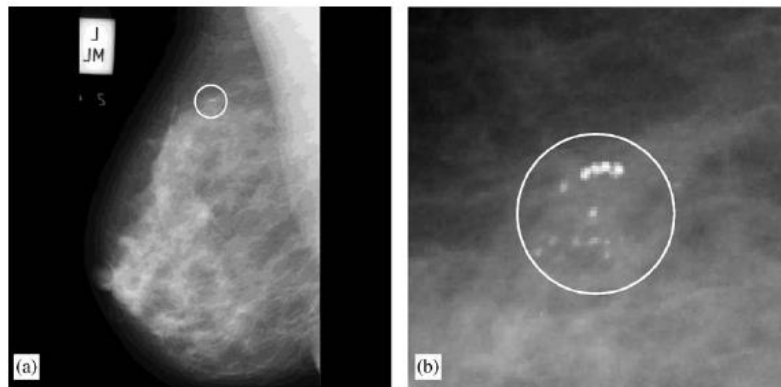


**Figure 5.9:** (a)  $\alpha$ -image using the maximum capacity measure. Microcalcifications are visible as little circles. (b) shows the results with minimum measure

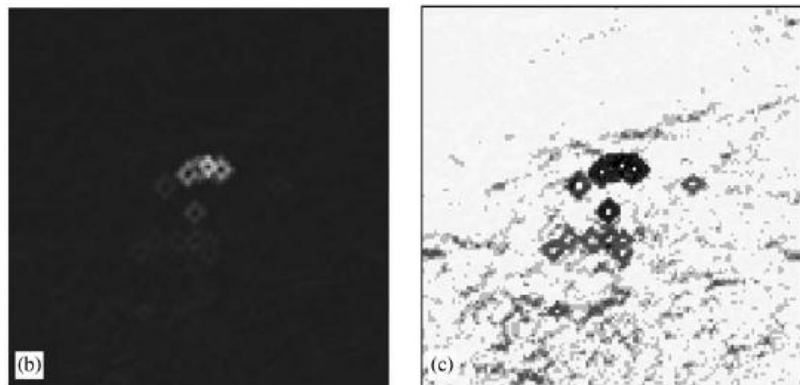
An additional decision is taken on the basis that microcalcifications are characteristically bright parts. It is already known that the calculation of  $\alpha$  relies on logarithm of measures. Logarithm acts as a magnifier on the lower gray band. Our interest is the opposite, since we would like to achieve the greatest resolution at high brightness levels. The adopted solution is straightforward: utilizing the inverted image, which forces to switch to the minimum measure capacity operator.

Usually, microcalcifications are areas with a relative high contrast and we can consider them as rare events. Such properties indicate where to look for them within the multifractal spectrum; pixels having high  $\alpha$  (high local contrast) and low  $f(\alpha)$  (rare events) will be the target. Just by selecting the pixels in a certain  $f(\alpha)$  range, and applying morphological operators in order to clean the result, this approach achieves really impressive results, even for hard cases. Next figures show the intermediate stages of the algorithm, and its final results.

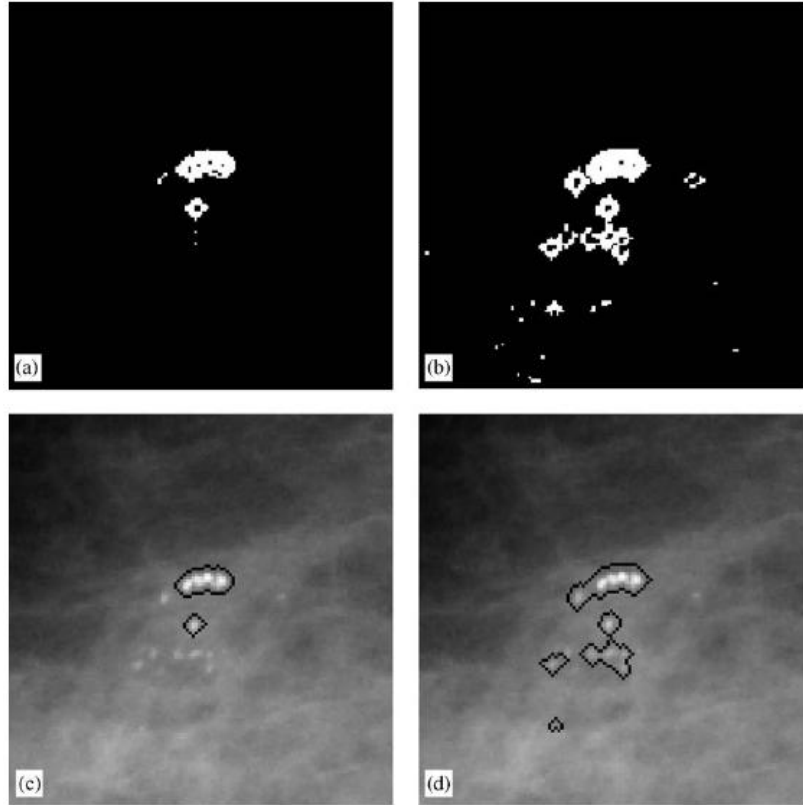
**Figure 5.10:** (a) Original mammogram (1024x1024) and (b) a detail view (128x128)



**Figure 5.11:** (b)  $\alpha$ -image. (c)  $f(\alpha)$ -image. White color corresponds to high levels. Notice that microcalcifications have a high  $\alpha$  and low  $f(\alpha)$ .



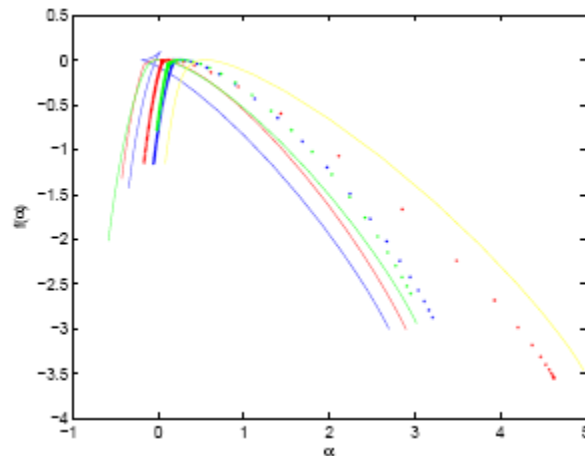
**Figure 5.12:** (a) mask formed by pixels having  $0 \leq f(\alpha) < 0.4$ . (b) mask for  $0 \leq f(\alpha) < 0.6$ . (c) and (d) represent the mammogram once the selection mask (a) or (b) are applied..



### 5.3.2.2 3D Classification algorithm

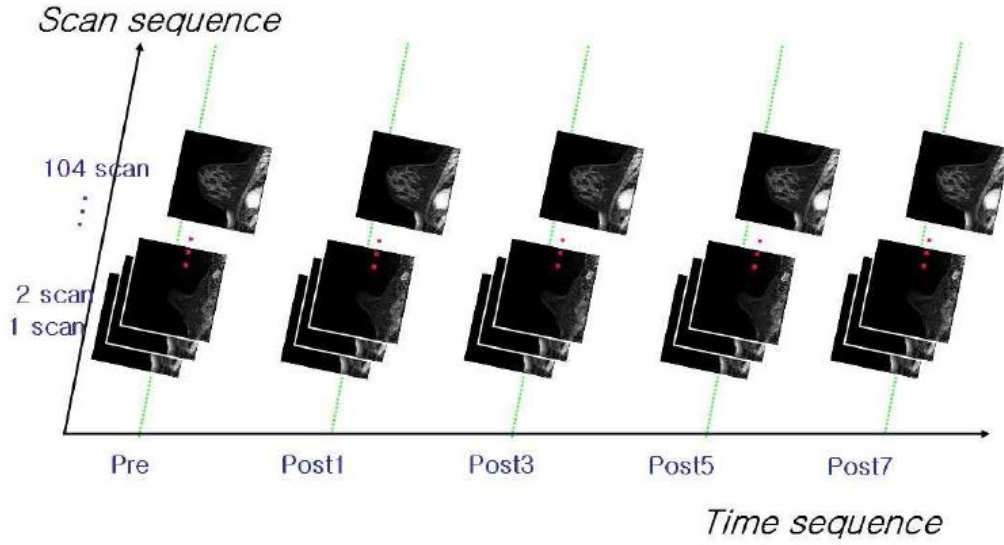
This section will describe the methodology introduced in [21]. The paper has been selected as a remarkable example of multifractal spectrum wavelet calculation. Additionally, and interesting contribution is achieved: the generalization of the concept of multifractal spectrum to the three dimensional case.

Authors define a framework for breast-area MRI image classification, with the aim of deciding whether a scan is either benign or malignant. For such a purpose, multifractal spectrum of the whole 3D data set is calculated by utilizing a DWT-based algorithm. The process obtains 7 spectrums, one for each DWT direction. Afterwards, in order to avoid dealing with cumbersome spectrum curves, some scalar descriptors are chosen to reduce the dimensions of the problem. Those descriptors will be in used in the image classification process.



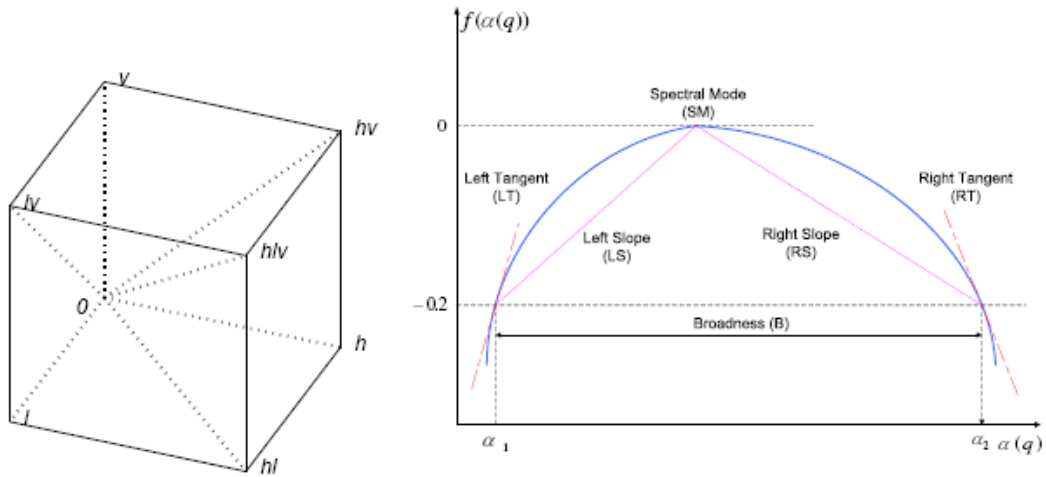
**Figure 5.13:** BRMI Multifractal spectrums for each direction

The data set utilized consist of time-series BMRI scans obtained from four women: 2 cancer and 2 cancer-free cases. The scan series includes one pre-contrast image and four post-contrast image acquired at 1, 3, 5, and 7 minutes after the contrast is administered. Scanning data set is formed by 104 sagittal slices by 256x256 pixels, with slice thickness of 3mm.



**Figure 5.14: Dataset representation**

Since Wavelet transform requires a cubic dataset, the 104x256x256 images are boundary mirrored in order to reach the 256x256x256 size. Afterwards, wavelet coefficients are obtained and a multifractal spectrum for each direction is calculated.



**Figure 5.15: Left: definitions of the directions of wavelet transform. Right: explanation of the characterization parameter of the multifractal spectrum**

Each directional-multifractal spectrum can be approximately described by 3 scalar descriptors without loss of the discriminant information, which are:

- 1) Spectral Mode (Hurst exponent,  $H$  or  $\alpha(0)$  ) represents the apex of spectrum or the Hurst exponent.  $H$  could be understood as a measure the global irregularity of a scan,
- 2) Left slope (LS) or left tangent (LT) represents the slope of the distribution produced by the collection of Hurst exponents with smaller values of the mode ( $H$ ). LS describe the

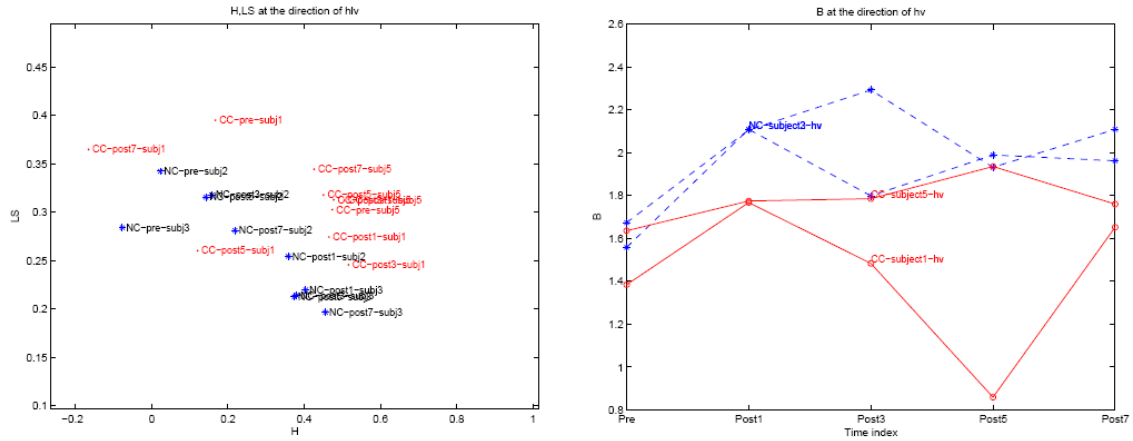
deviation from mono-fractality. The higher LS values, the greater mono-fractality. Theoretic pure fractals (monofractal) have infinite LS.

- 3) Width spread (Broadness, B) or right slope (RS) or right tangent (RT). Broadness allows easily calculating RS and RT, while the reverse process is not so straight forward. Hence, B is the preferred over other parameters. B could be seen as a summary of the overall spectrum, showing the variability of the Hurst exponents.

In the paper, BMRI data available consists in two cases and two controls.

For all practical purposes, the descriptors which provide more accurate classification results are H and LS. In particular, the value is the one corresponding to  $h/v$  direction, which is the one that has pure high-frequency data. Given that set of markers, it is stated that healthy patients show a smaller H+LS value than cancer affected. Regarding the temporal behavior, it seems that studying the evolution of parameter B in  $h/v$  direction shows significant selection capabilities.

The authors stated that their findings are consistent with empirical evidence that healthy responses are characterized by irregularity (multifractality and low H+LS) and that increased regularity (monofractality and high H+LS) may suggest pathologies.



**Figure 5.16: Left: LS vs H(h/v direction) case classification. Right: temporal evolution of B. CC: Cancer Cases. NC: No Cancer**

Due the limited range of images studied, results are not conclusive and the applicability of formal classification is limited. In addition, the selection of decision parameters and thresholds is not justified. In conclusion, further study is required, but it is an interesting starting point, giving some clues for future developments.

## 6 OVERVIEW OF THE APPLICATION FOR HEPATIC TUMORAL TISSUE DETECTION

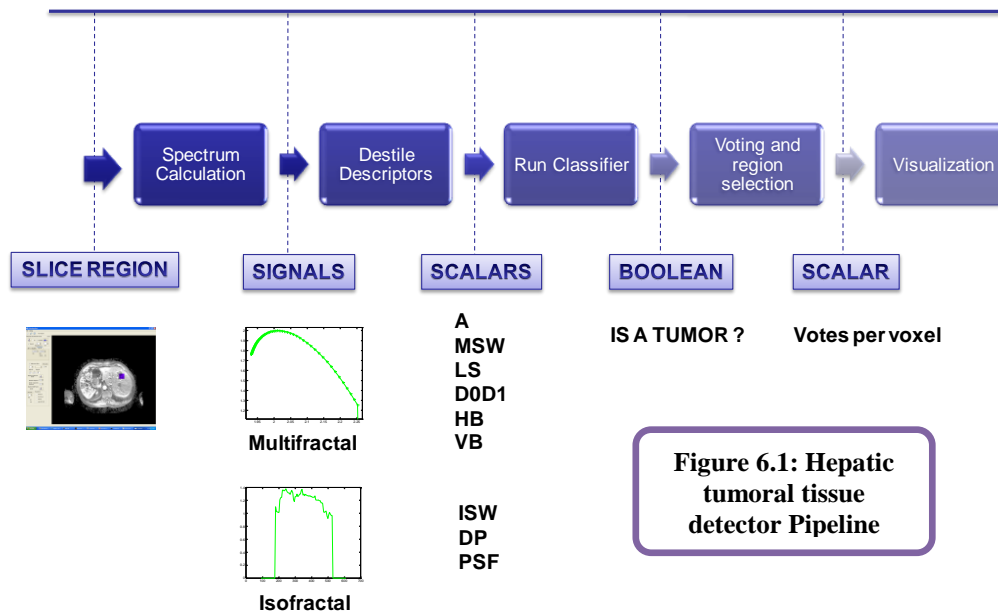
This chapter is the starting point for the second block of the thesis, mainly presenting the development of a piece of software aimed to the detection of metastatic tissues in the liver through the usage of fractal analysis.

The application has been created having in mind two main purposes. In the first phase of research, it was utilized as a tool for the application of fractal analysis techniques in the regions of interest and studying its behaviour. In a second phase, and in consonance with the thesis main goal, the application was extended to be able to detect and highlight areas containing tumoral tissues in the liver.

Figure 6.1 depicts the pipeline of the detection algorithm that has been implemented. The process traverses the whole image volume, using an analysis unit composed by a 32x32 planar intraslice window. The window is slid in order to cover all the possible positions in the volume. For each window location, the system performs two type of Fractal analysis, obtaining the corresponding Multifractal spectrum and the Isofractal Spectrum, which is a new Fractal technique developed in this thesis. Chapter 7 fully covers the theoretical details and implementation issues regarding those fractal techniques.

In the following stage, the obtained spectrums have to be summarized into a few meaningful scalar values in order to ease their classification. That is the role of the descriptors. Once the descriptors have been calculated, it is necessary to select which of them are more significant for tumor detection, and to establish the thresholds that separate benign from malignant tissues. Both decisions, referred as characterization phase, lead to the construction of tissue classifiers. The process of characterization of the spectrums, the distillation of descriptors and the classifier construction is detailed in the Chapter 8.

Finally, the classifier decision leads into a windowed region tagged as malignant or benignant. In the former case, the voxels within the window receive a vote. Once the entire set of windows has been analyzed and the voxels have their vote results, the visualization of the model is performed. In this stage, the idea is to highlight the voxels using the number of votes as the selection intensity, allowing an enhanced visualization of tissues with higher probability of malignity.

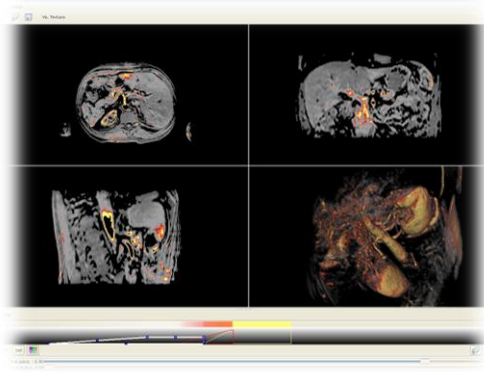


**Figure 6.1: Hepatic tumoral tissue detector Pipeline**

## 6.1 System Overview

All the software developments concerning the thesis has been integrated as a module in the Medical Imaging platform *VRMedVolVisDicom*, created by the MOVING<sup>7</sup> research group from Universitat Politècnica de Catalunya (UPC). The underlying programming language for all the works has been C++

VRMed platform provides facilities for DICOM format management, easing the tasks of image and metadata loading, which is crucial since most of the medical images obtained from imaging equipments are compliant with the extensive *DICOM* standard (*Digital Imaging and Communications in Medicine*).

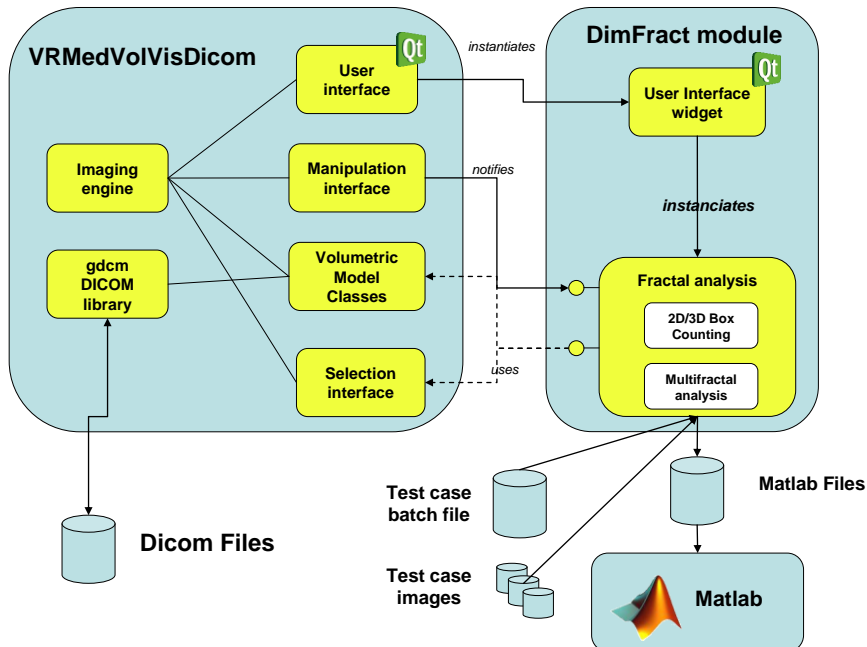


**Figure 6.2: A screenshot of the VRMedVolVisDicom basic interface**

The developed module, from now on *DimFract*, utilizes VRMed visualization capabilities to integrate analysis results into planar volume sections and volumetric views of the model. The application also generates some statistical and plot results in the form of Matlab<sup>8</sup> language and stored in plain ASCII files. Matlab facilitates graph generation and allows any kind of data manipulation for further data exploration.

The graphic user interface has been designed following the design guidelines defined by the VRMed team, based on the MVC pattern. The GUI development environment is based on the multiplatform Qt 4.4.3 library and QtDesigner from TrollTech.<sup>9</sup>

Summarizing, Figure 6.3 shows the overall architecture of the system, delimiting the VRMed and DimFract interfaces, as well as its interaction with external data files.



**Figure 6.3: System architecture overview**

<sup>7</sup> <http://www.lsi.upc.edu/~moving/>

<sup>8</sup> <http://www.mathworks.com/>

<sup>9</sup> <http://www.qtsoftware.com/>

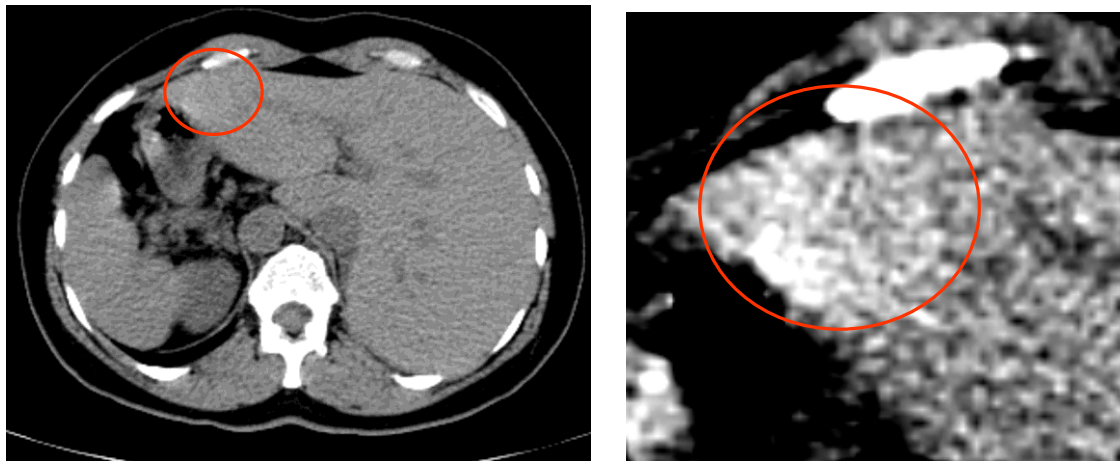


## 6.2 Data sets

DICOM data sets utilized in this project were provided by the *Vall d'Hebron* Hospital, which include several confirmed cases of hepatic metastasis. Since the information for each patient was heterogeneous, one of the first tasks was to select the study cases amongst all available data, finally deriving in two representative cases, named **A** and **B**.

	Case A	Case B
Type:	High resolution MRI. T1-Weighted / FAME (Fast Acquisition Multiphase Efgre3D) Contrast protocol, 4 MRI were taken: Basal-Arterial-Venous-Retarded.	Low resolution CT scan.
Range:	Liver-centered	Full head+torax+abdomen
Image size:	512 x 512 x 76 voxels	512 x 512 x 257 voxels
Dimensions:	460 x 460 x 190 mm	500 x 500 x 840.4 mm
Bytes/pixel:	2	2
Voxel size:	0.898438 x 0.898438 x 2.5 mm	0.976562 x 0.976562 x 3.27 mm

The start-up idea of the project was to search for variations of fractal dimension provoked by tumoral angiogenesis. Hence, it was necessary to have images with a high level of definition in the affected areas, good enough to provide details due to vessels presence. One of the first impressions was to notice that normal CT images provide a poor accuracy in liver soft tissues, and they do not allow appreciating vessel structures. Figure 6.4 shows case **B** CT scan, where the liver and a metastasis area are shown. A detailed view is provided, where the gray levels scale has been modified ad-hoc to highlight the affected area. As it could be appreciated, the information on standard CT is on the texture (or noise) and the gray levels, rather than in the shapes.



**Figure 6.4:** CT slice corresponding to case A

In contrast, MR provides an enhanced view of the liver area, which allows differentiating tissues and vessels. Case **A** contains is a full MRI protocol, which is a complete studio consisting in different scan series varying the caption parameters. The studio includes T1 and T2 weighted images. For our purposes, T1WI images are more adequate, since they detail the vessel areas better. Furthermore, T1 images include a contrast time series, which permit temporal series analysis.

## 7 DEFINITION AND IMPLEMENTATION OF FRACTAL ANALYSIS ALGORITHMS FOR TUMOR DETECTION

This chapter covers the definition and implementation details of the fractal analysis tools used for determining whether the tissues are healthy or malignant. Two main algorithms have been developed, the Multifractal Spectrum, calculated through the method of moments, and the Isofractal Spectrum which is one of the main contributions of this master thesis. Both techniques study the image set from different points of view; multifractal focus on the fractal distribution of irregularity, while Isofractal Spectrum studies the fractal geometry of isosurfaces.

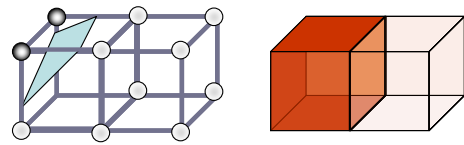
In initial phases of development, the primer input of ideas was the definition of Isosurface Box Counting, which consists of the calculation of the fractal dimension of an isosurface, i.e a surface that links images areas sharing the same density value. This technique was not sufficient, and had to be enhanced, leading to the creation of the Isofractal Spectrum, which could be understood as the representation of the fractal dimensions of the isosurfaces for the entire range of densities.

The chapter finally describes how those fractal analysis tools could be applied into the images using the application and the way its results could be visualized.

### 7.1 Isofractal spectrum analysis

#### 7.1.1 Isosurface Box counting

If we consider the volumetric model as a 3D set, voxels could be tough as the nodes of a 3D grid. Each node has an associated value: its density value. If we choose a certain density value threshold, it is possible to extract a surface that separates the volume where the density value is greater from the volume where the density is lower than the threshold. This is called an isosurface. Multiple algorithms have been designed to extract an isosurface, most of them based on the marching cubes (MC) algorithm.



**Figure 7.1: (left) grid of black and white voxels. (right) equivalent box counting view. Left box is countable, right box is not countable.**

The approach presented in this work consists of calculating the fractal dimension of the isosurface using boxcounting. Since we are not interested in the surface itself, but in its FD, it is not necessary to explicitly calculate the mesh surface. In order to apply 3D box counting, it is enough to be able to determine whether a cubic grid box formed by 8 voxels is crossed by the isosurface. If this is the case, the box is countable. The calculation is simple, for each voxel grid, determine whether is black (its value is over the threshold) or white (below the threshold); the grid box is crossed by the isosurface if there exist at minimum one white and one black vertex. The same process could be defined in 2D, just by working on a single scan slice. The process remains basically the same, but working with 2D squares, and isocurves.



**Figure 7.2: (left) 3D surface obtained from an image, interpreted as height-map. (right) an iso-curve extracted from the height-map**

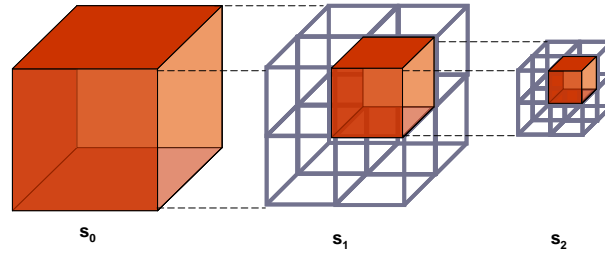
The initial study was aimed to determine whether this kind of analysis could provide significant information for tumor location.

At this point, it is interesting to recall differential box counting (DBC). DBC represents a 2D image as a height map which defines a 3D surface, in order to apply 3D box counting. In contrast, IBC calculates the FD of the iso-height curves, instead of calculating the FD of the height-map derived surface (Figure 7.2).

### 7.1.1.1 Implementation

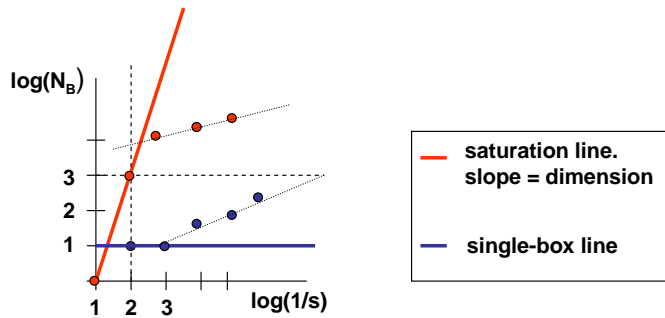
The IBC has been developed both in 2D and 3D versions. From implementation point of view, the calculation of the box counting has been designed as a recursive process. The recursion starts at the lowest scale level ( $s_0=1$ ), from a box containing the entire analyzed volume, whose size is a power of two. At each recursion step it is needed to decide whether the current box at scale  $s_n$  is countable, which in turn requires dividing the box using the next scale level ( $s_{n+1} = s_n * 2$ ), and asking whether any of the smaller box contained is countable. If true, the current box is countable.

The trivial case of recursion happens when the maximum scale is reached, the box coincides with the unit-box. Here, just looking at the threshold and voxel values, we can decide if the box is countable.



**Figure 7.3: Different stages of box counting recursion**

In order to accelerate the unit-box case, a cache has been created for storing the maximum and minimum density values of the voxels belonging to each unit size box. If the threshold value is in the range between max and min, the box is countable.



**Figure 7.4: Linear regression on degenerate cases in a 3D box counting.**

Once the recursion is completed, the count of boxes ( $N$ ) for each scale level is collected, and using linear regression, the slope of the plot  $\log(N)$  versus  $\log(1/s)$  is calculated. In order to obtain an accurate slope, the algorithm selects the samples that need to be used in regression. Two degenerate situations are considered: saturated areas, and almost-empty areas. Saturated areas are those containing a high number of active boxes, and all the boxes at low scale levels are countable. Till a

certain scale level is reached, non-countable boxes do not appear.

In the other hand, almost empty areas suffer from the opposite effect, for small scales, only one box is active. This is constant until enough zoom is achieved, and multiple active boxes appear.

Both effects are shown in Figure 7.4, the algorithm simply discards the consecutive samples that lie on the saturation line defined by:

$$y = \text{dimension} \cdot (x-1)$$

or in the single-box line  $y=1$ . It permits obtaining a most accurate slope, since the discarded samples were not meaningful due to the lack of resolution.

### 7.1.1.2 Scale range

The nature of iterative implementation enforces a geometric evolution of scales, the resolution is doubled at each step. Hence, given a study area of size  $32 \times 32 \times 32$ , the box sizes applied are 32, 16, 8, 4, 2, 1. The number of scale levels is directly obtained from  $N_{levels} = \log(\max\_size) + 1$ .

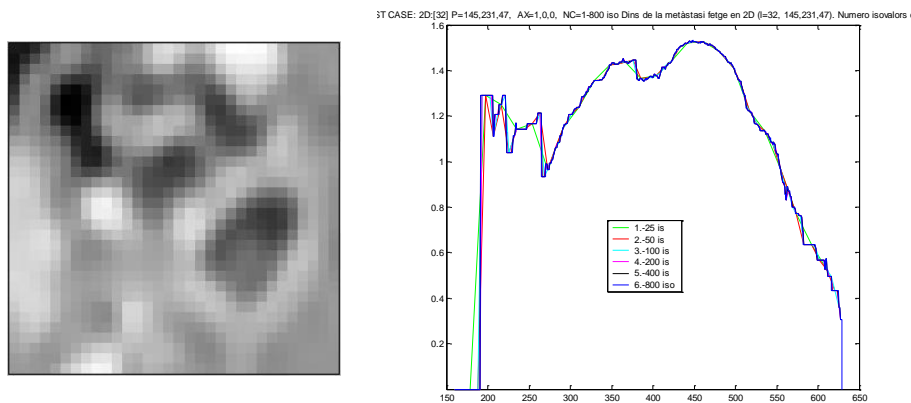
Alternatively, an iterative solution has been tested. This solution is able to provide any set of scales, at the cost of increasing the computation time due to the augment of scale levels and the lack of reutilization of the high scale information in the low scale data which is naturally provided by the recursive solution.

The iterative option has been tested, using box-side sizes ranging from 1 to total area width in unitary increments, but the results achieved in the tests were not satisfactory when compared with the recursive choice.

## 7.1.2 Isofractal spectrum

After preliminary tests on the medical images, it came to the light that the selection of a threshold isovalue is a burden, which enforces human interaction and multiple trial and error. In some cases it is possible to find a certain value which shows an increase of fractal dimension in the areas of interest, but results are not generalizable. The situation is worsened by the fact that MRI do not guarantee scaled density values, which means that if some interesting thresholds are found, they will be not reusable for other similar images since the density scale will be different.

In the other hand, it is quite probable that the information is not tied to a given density value, but to the interaction between density levels within the data set. With that in mind, it becomes natural to think of studying the evolution of fractal dimension along the density values existing in an image, which here it is named after Isofractal Spectrum (IS).

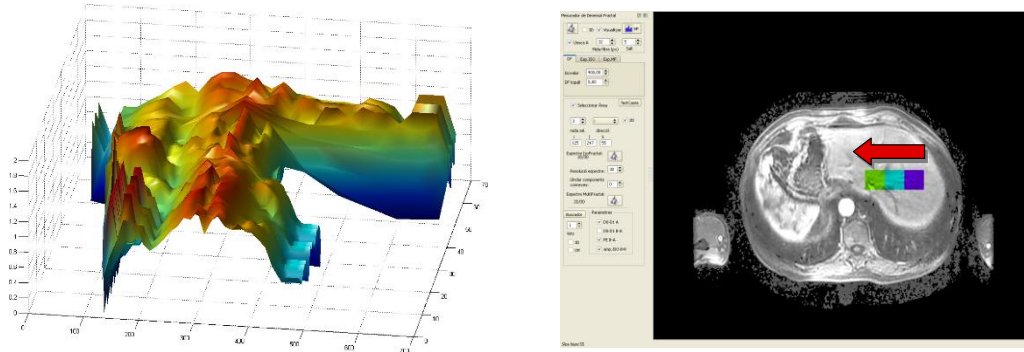


**Figure 7.5:** grey scale image obtained from CT (right), and its corresponding isofractal spectrum (right). Series correspond to different resolution of isolevels.

Given the analysis set, the maximal and minimal density values are obtained. The range is then homogeneously quantized in  $R$  segments, and IBC is calculated for the resulting  $R+1$  sampled density values.  $R$  becomes one of the critical parameters defining the spectrum, since in one hand determines the resolution of the spectrum, and in the other regulates the computational effort.

### 7.1.2.1 The scanline spectrum

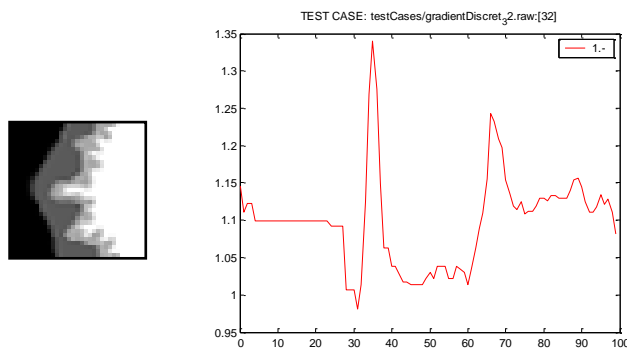
In order to study the behavior of the iso fractal spectrum in regions of interest, DimFract module also provides of a functionality for representing the evolution of the IS along a scanline. By selecting a direction aligned to one of the axis ( i, j or k ), a starting point and a path length, the evolution of the spectrum could be represented graphically as a 3D graph, as shown in the Figure 7.6.



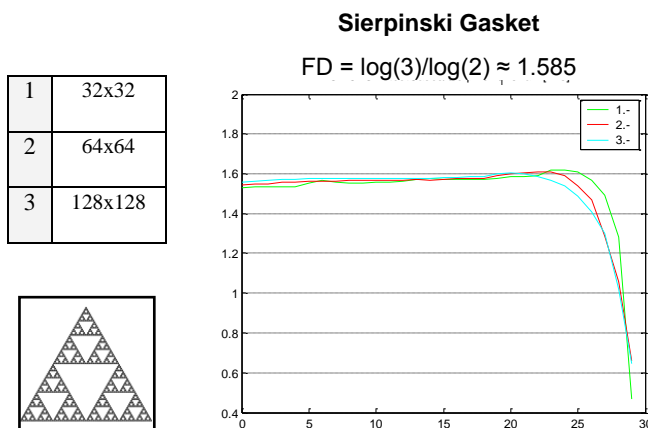
**Figure 7.6:** (left) scanline spectrum corresponding to the scanline defined by the user in the application (right).

This kind of representation may be helpful when determining the changes that characterize the transition from an area of interest to another.

### 7.1.2.2 Significant test cases



**Figure 7.7:** Case TC.A: IS for discrete gradient

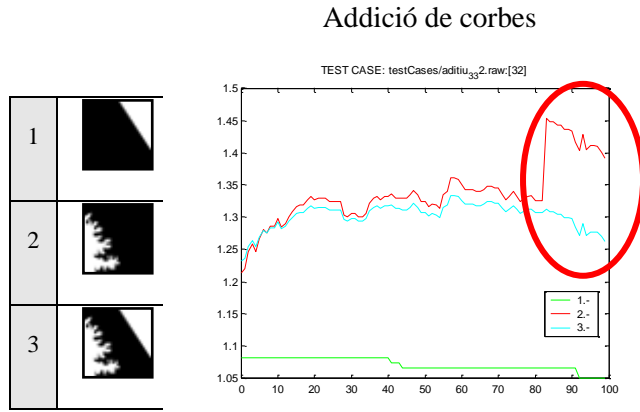


**Figure 7.8:** Case TC.B: IS for sierpinsky gasket.

A large set of test-case images has been run against the IBC algorithm to learn about its behavior in different situations. Left column shows two cases where the results are according to the expectations.

In TC.A there are 2 meaningful density values, showing increasing irregularity. This is perfectly depicted in its fractal spectrum, where the two peaks show a remarkable increase from reference dimension (1- straight line).

Additionally, the algorithm has been run on a deterministic fractal, the Sierpinski gasket. As can be seen in case TC.B, the approximation to the exact value of fractal dimension (1.585) is pretty good. It may be surprising that the spectrum does not show a single peak of height 1.58, this is provoked by the antialiasing of the image, which creates density gradients around the



**Figure 7.9: Case TC.C: IS evolution for curve addition.**

Theoretical results state that in the addition of sets, the fractal dimension should be that of the set having greatest FD. As could be seen in series 3, there is a sudden peak coinciding with the curve threshold. The cause of this non-desired effect is the insufficient level of resolution, which contaminates the slope calculated with regression. This anomaly remains as an open problem.

### 7.1.2.3 Minimum Connected-Component filter

Isosurface spectrum goes through different density values and, for each of them, an isocurve is extracted. In fact, the isocurve is usually formed by a set of disconnected smaller isocurves. If we consider the effect of random noise in the images, it is easy to see that it introduces an undetermined number of small disconnected isocurves at different density values. These microcurves may alter the calculation of the regression, whose negative effects have been already noticed in the case TC.C. Hence, these noise-derived isocurves may alter the values of fractal dimension of the target curves. Since the goal pursued in this work is to find traces of the vessel structure, the main interest is focusing in isocurves with a noticeable connected length. Therefore, an extra preprocess step has been added to the IBC in order to get rid of small curves: a minimum connected-component filter.

During the spectrum calculation, the filter is run once for each sampled density value. An image is formed containing all the isocurves corresponding to the current sampled density level. Here, the filter finds the connected components and labels each of them with its length. The algorithm works on a recursive way, using a seed-propagation strategy. Later on, the connected components which do not fulfill a minimum length requirement are discarded. Finally, FD is calculated on this reduced set.

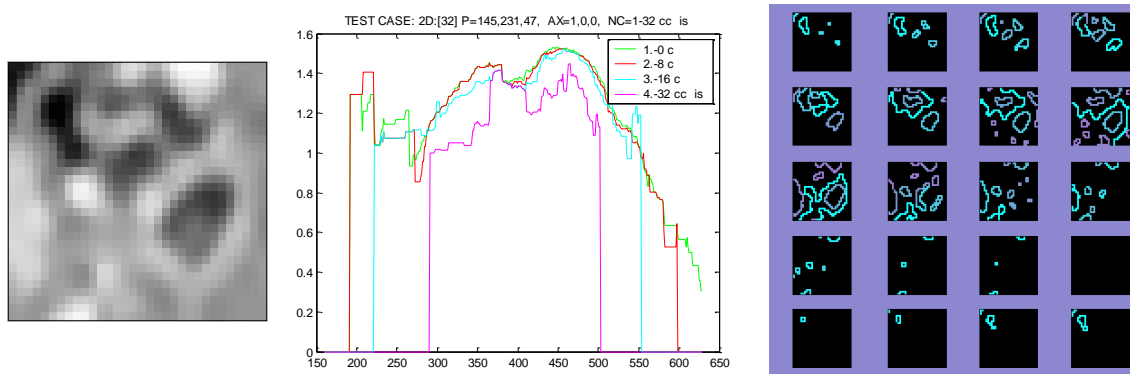
Hence, a new parameter is required for IBC: the minimum connected-component curve size ( $cc$ ). From the implementation point of view, once the connected components are labeled, during the box counting process all the areas that belong to curves with a size below the  $cc$  threshold become not countable.

The application of this filter eliminates spurious areas of the spectrum while maintaining the most significant peaks. Figure 7.10 shows the effect of increasing the minimum connected size from 0 to 32 units in the spectrum calculation.

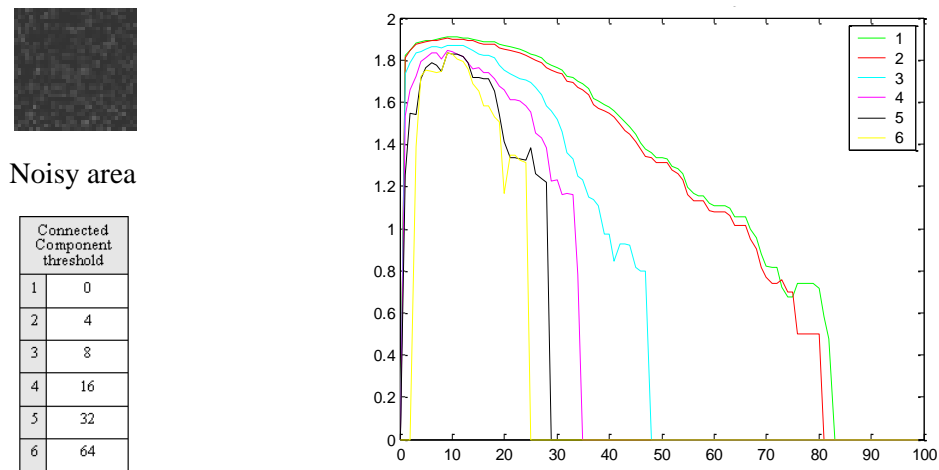
sharp edges. In any case, the gradients have a similar dimension to the set itself.

TC.C introduces one of the degenerated cases that put some constraints to the reliability of this technique. Spectrums for Images 1 (straight line:  $DF=1$ ) and 2 (complex line:  $DF>1$ ), work as expected. The problem appears when adding both curves.





**Figure 7.10: Effect of the connected-component filter in Isofractal Spectrum. (right) isocurves for different density values. Connected components are labeled with a color code which represents its length.**



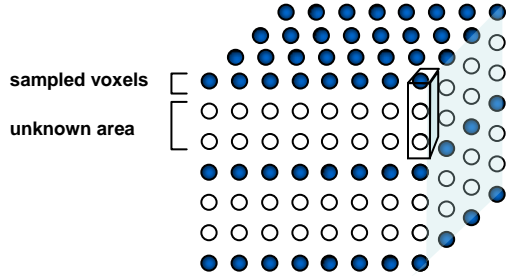
**Figure 7.11: Isofractal spectrum of a gaussian noise affected area, and the evolution when increasing the connected components filter threshold.**

#### 7.1.2.4 Regression coefficient

Linear regression provides a quality indicator named regression coefficient, so-called Pearson coefficient. This value is proportional to the summed squared residuals (the difference between the sample and its linear approximation), and it is a good sign for determining the applicability of the linearization. The initial guess was to give an important weight to this parameter, in order to discard spectrum samples which not guarantee a minimal regression coefficient, but it has been stated that most of the samples its value is above 0.89. The cause of that phenomenon is the scarce available resolution, whose direct consequence is that there are only a little samples for slope calculation. The situation is worsened by the effect of saturation and single-box areas, which finally derive in a further reduction of the number of samples. The linearization using small number of samples usually generates low error rates.

### 7.1.3 Problems with resolution in volumetric models

One of the most remarkable characteristics of volumetric images is their limited resolution in the axial direction. This is especially dramatic for ordinary CT images, which supply the poorest resolution. In our data sets, the proportion of voxels' interslice size (i,j) dimension versus intraslice size (k) is 2.78 for MRI and 3.35 for CT. Hence, the spatial sampling anisotropy is around a 3 factor, which greatly questions the validity of exploiting the spatial coherence in the k axis direction.



**Figure 7.12: Approximate representation of the volumetric model data.**

The three-dimensional IBC requires working with cubic windows, having equal number of samples in each direction. In this case, it means that in the real space, the “cube” is in fact a right prism of square, whose height is 3 times the base side. For increasing window sizes, the effect of anisotropy is devastating, in most of the cases is not possible to apply the fractal analysis to the desired areas, since the range in k direction is so large that the studied area becomes highly heterogeneous, which in turn invalidates the calculation of fractal dimension.

A naïve work-around would be using trilinear interpolation for providing estimated samples in order to recover isotropy. Again, this approach is not attractive, since we are speaking of generating data for a 3mm gap between slices. The structures that we are looking for could fit completely inside this intraslice space; small vessels' branchings, local variations of tissue density, holes, and other structures may fall fully between slices.

The resolution limitation finally leads to a disfiguration of the results obtained using 3D analysis windows, unless tiny areas were used, which will not guarantee a precise fractal dimension calculation. Both iso fractal and multifractal spectrum suffer the effects of this problem.

## 7.2 Multifractal spectrum analysis

Iso fractal spectrum is related to a geometrical interpretation of the set, where the isocurves are studied and their fractal dimension calculated. Multifractal approach is much more abstract, here we will be calculating the fractal dimension of areas with the same singularity behavior, which has nothing to do with the density level, but with the variation of density levels in different scales. It could be thought that the information obtained from multifractal analysis is complementary to that calculated with iso fractal spectrum.

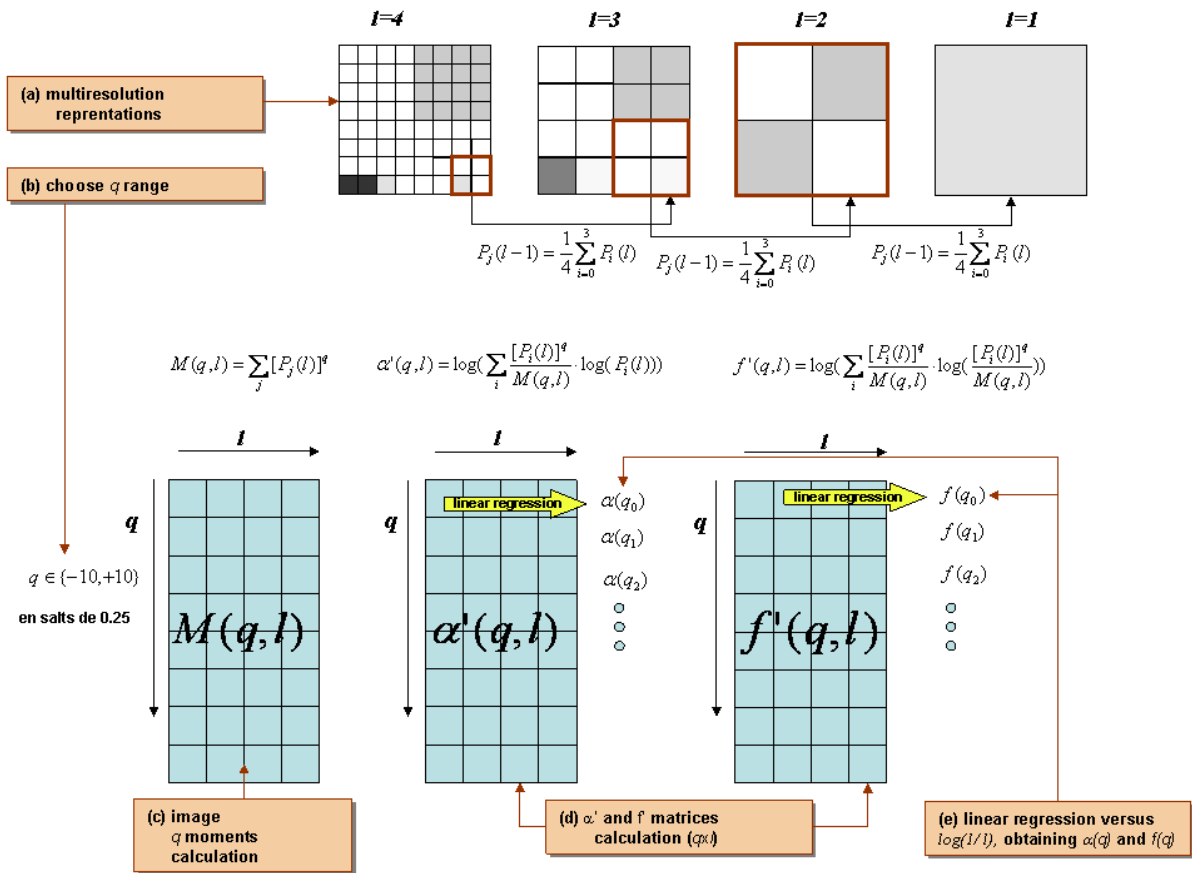
The implementation of multifractal analysis is based on the system introduced in [9] by Chhabra and Jensen, which is defined by the equations (4.11)(4.12)(4.13). The method requires computing the moment  $q$  of the measures at several scales. The calculation uses a normalized measure, which help avoiding numerical overflows, as defined in (4.11). The voxel's density value is directly used as cell probability  $P$ .

$$(4.11). \quad \mu_i(q, l) = \frac{[P_i(l)]^q}{\sum_j [P_j(l)]^q}.$$



The algorithm requires the following steps:

- A multiscale representation is calculated from the original voxel set. Scale is indicated as  $l$
- Choose a set of  $q$  values for which the Multifractal spectrum is going to be calculated. In this case the selected range is  $q \in \{-10, \dots, +10\}$  using steps of size 0.25.
- For each scaled image and  $q$  value, calculate the matrix of total moments  $M(q, l) = \sum_j [P_j(l)]^q$
- For each scaled image and  $q$  value, calculate the matrix  $\alpha'(q, l)$  and  $f'(q, l)$
- Apply linear regression to the each of the rows of  $\alpha'(q, l)$  and  $f'(q, l)$  against  $\log(1/l)$ . This will provide  $\alpha(q)$  and  $f(q)$ .

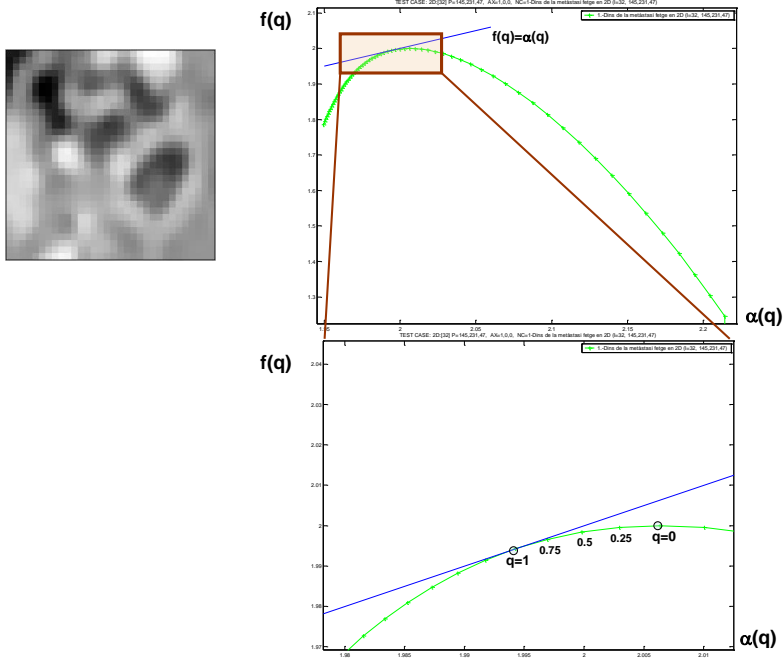


The implementation is quite similar in 2D and 3D, the main differences are found in step (a), where in the three-dimensional case, the sum of 4 voxels forming a 2x2 square is replaced by the sum of 8 voxel defining a 2x2x2 cube.

In some cases the linear regression will not provide a reliable slope because of the samples distribution is not linear. Those cases where the coefficient of regression is below 0.97 are discarded in the final spectrum. That usually happens for values of  $q$  situated in the low and in the high band of the range.

### 7.2.1 Validation

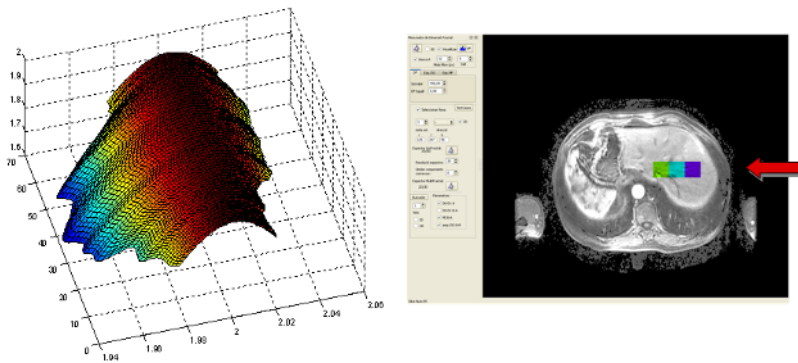
The Chhabra and Jensen method provides an accurate spectrum, especially in the area near  $q=0$ . In order to verify the correctness of the algorithm, it has been verified that the resulting spectrums are compliant with the basic properties of the multifractal spectrum. As could be observed in Figure 7.13, the whole curve is below the bisector  $\alpha(q) = f(q)$ , and  $f(q) < f(0) = D_0$  for any  $q$ . Furthermore, the curve is tangent to the bisector at  $q=1$ . The value of  $f(1)$  is often referred as the information dimension or  $D_I$ .



**Figure 7.13:** (top right) Multifractal spectrum of a MRI 32x32 area (top left), and a detail view (bottom) showing the tangency with the bisector at  $q=1$ .

### 7.2.2 The MF spectrum scanline

In a similar manner, the variation of the singularity spectrum along a scan line defined within the image could be analyzed. In this analysis, the variation of the spectrum broadness is easily visualized, but the slope or  $\alpha(0)$  (maximum positions) variations are difficult to perceive since the scale of variation is really small when compared with the whole spectrum's broadness.



**Figure 7.14:** Multifractal spectrum scanline

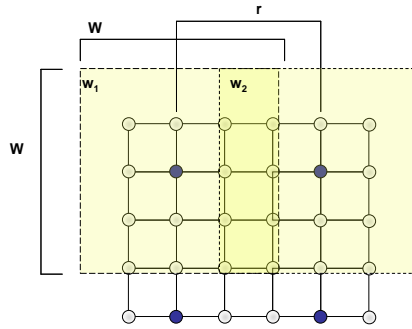
## 7.3 Application of the fractal analysis techniques

The DimFract module is a workbench which offers to the user different alternatives of application of the fractal algorithms which has been presented to the volumetric models. Those processes are well classified by using the following ontology:

- **Global or local:** differentiating whether the application of the process involves the whole image or just affects a restricted area.
- **Static or dynamic:** regarding the temporal applicability. Static methods work in a single image, while dynamic work in multiple temporal snapshots.

### 7.3.1 Global operations

Global operations are related to the application of a fractal analysis operator to the volumetric model completely, or in a whole slice otherwise. Such operator must be able to obtain a scalar value from a small window of size  $W$  centered on a given voxel. The operator is run for every voxel found on a regular grid of size  $r$ , which is equivalent to slide the window through the image in gaps of size  $r$ . If the maximum resolution is required,  $r$  is set to 1, which means that the calculation is performed for all the voxels, but for the sake of running time, it is reasonable to adjust  $r$  to greater values (3, 5...).



The selection of an adequate value for  $W$  is high dependant of the application context.  $W$  regulates the accuracy of the calculation (the greater, the most) and the area precision (the smaller, the best), both aspects are important and we need to find a balance between them. A good heuristics consist of choosing the greatest  $W$  which ensures resolution enough when compared with the target ROI size. Finally, there is another important issue, big values of  $W$  increase the process time, but this could be compensated using greater  $r$ .

**Figure 7.15: Calculation of the Fractal image**

Once the scalar values have been calculated, they are visualized in combination with the model. The system merges the luminosity of the original density image with the hue obtained from a color mapping of the scalar value generated by operator, based on HSV palette. If  $r$  is greater than 1, the scalar value for a given voxel grid is extended to its neighboring voxels (using  $r/2$  radius) where the operator result is not available.

But, which are the available operators? In fact, there are multiple of them, belonging to one of the three categories: iso fractal dimension, iso fractal spectrum and multifractal spectrum. With exception of the IFD, it is necessary to define parameters that summarize the spectrums information in a single scalar. The parameterization is covered in the following sections.

Next figure is an example of a fractal image representing the fractal dimension for the density 1090, which delivers a proper segmentation of liver zone. The image has been obtained using  $r=3$  and  $W=16$ .

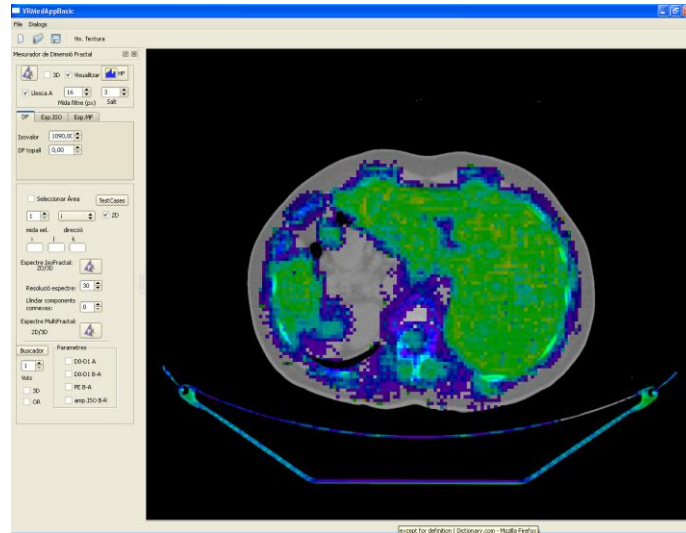


Figure 7.16: Iso-fractal Dimension image for the density value 1090.

### 7.3.2 Local operations

In some situations, it is required to study specific regions of interest. DimFract permits selecting interactively those areas and obtaining their characteristics multifractal and iso fractal spectrums, given a set of input parameters. The raw data and the plot are written in a Matlab file, which could be directly executed within the Matlab GUI.

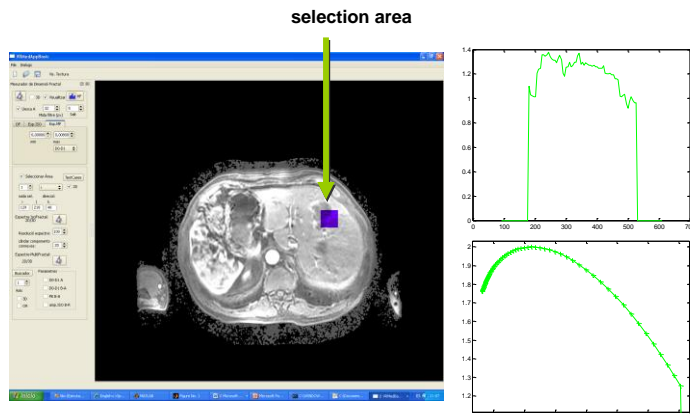


Figure 7.17: manual selection of ROI and iso fractal (top) and multifractal (above) spectrums calculation.

The application is also capable to batch the analysis processes through the usage of a test configuration file. The file contains a list of operations to be executed; each of them specifies the 3D coordinates of the areas to be analyzed and the type of analysis to be performed as well as the input arguments for each process. It is also possible to create series of operations that will be plotted together in order to ease the comparisons. This capability has been really helpful in the characterization of ROIs. For more details on the file

structure, please refer to the appendix 13.6.

Finally, the last local operation is the scan-line analysis. This functionality demands the user to select a region of interest, the axis direction and length of the scan line where the spectrum is going to be calculated. This time, the results come into the form of 3D plot, since we will be visualizing the spatial evolution of the fractal spectrum.

### 7.3.3 Dynamic operations

We refer to dynamic operations when the data utilized is coming from multiple images of the same patient captured in different instants. Precisely, the working data set A is a contrast MRI formed by the basal image (before the contrast) plus three images (namely, arterial, venous and retarded) which show the diffusion of the contrast agent propagated through the vessels into the tissues.

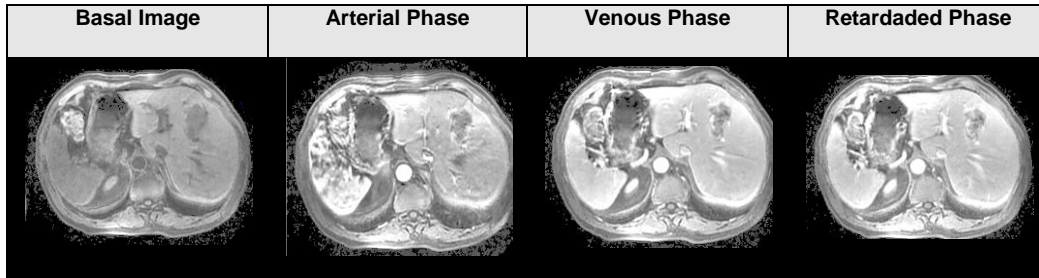


Figure 7.18: MRI contrast image sequence

Since the VRMed application has been designed to work with an image at a time, the management of dynamic operations is provided in a rudimentary manner. The four images have been preprocessed using fractal analysis and the main descriptors have been saved in independent custom files in order to be able to load the required information from multiple sources from the software.

## 7.4 Summary

In this chapter, the fractal analysis techniques utilized in the project has been detailed. The novel Isofractal Spectrum gives a geometric vision of the window, summarizing the fractal dimension of all the isosurfaces within a region. In its calculation several issues have arisen, specially regarding to the lack of resolution and its effects on the regression, which were amended by discarding degenerated samples.

Another problem detected in the calculation of Isofractal Spectrum was the response to additive combination of signals. Due the lack of resolution, the result differs from theoretical expectations, since the regression lines are contaminated with the contribution of different signals, which provoke fractal dimensions to be higher than reference values. Noise has similar effects on the images. In order to reduce the effect of small isocurves, the usage of a connected component filer has been proposed. It allows focusing in the contribution of the largest isosurfaces, lessening the weight of noise of low FD areas.

The fractal analysis could be done in 3D or 2D regions. It has been shown that the sampling anisotropy of the volumetric images do not permits to utilize the 3D option, since the size of the region and the lack of spatial coherence in the axial direction invalidate the obtained results.

Regarding Multifractal Spectrum, the implementation has been done using the method of moments and its variant for direct  $f(\alpha)$  calculation, which permits a light-weighted processing while obtaining an accurate spectrum.

The chapter is closed by the classification of the application scopes of the fractal tools. It could be summarized as local versus global application (a single delimited region versus the entire volume), and static versus dynamic (within a single image or using multiple images in a timeline).

## 8 ANALYSIS AND CLASSIFICATION OF MRI BASED ON FRACTAL CHARACTERIZATION

At this point, we are aware of the analysis tools that have been packed within *DimFract* module. The second part of the project, and probably the more time consuming, has been the study, definition and test of tissue classifiers. The complexity of this phase is related to the high number of input parameters and to the fact that the algorithms output are spectrums, which is to say signals. That point makes the problem rather more uncomfortable to manage. In this context, the best strategy applicable is to search for scalar descriptors of the fractal spectrums in a way that we retain the characteristics that allow differentiating tumoral tissues from healthy tissues. Lately, the ranges of descriptors which correspond to malignant areas will be set, and classifiers will be obtained by the combination of multiple descriptors.

The effort on that stage has been centered in the data set A, the MRI with contrast, which is formed by four volumetric images: basal, arterial, venous and retarded phase. For static analysis the arterial image has been used since it provides the best vessel representation. Dynamic analysis involved all the images.

It is fair to advance that after the primer test, the 3D analysis was discarded, since the results obtained were not applicable to tissue classification. This decision has his foundations in all the problems already detailed in previous chapters. Therefore, all the results provided here will be dealing with 2D intraslice analysis.

### 8.1 Methodology

For static characterization, a test set formed by 35 MRI regions of 32x32 voxels have been defined, which will be referred as Regions of Interest (ROI). These regions are representative for **non-liver tissues** (10 samples distributed in different types of tissue: pulmonary, knees, peritoneal areas, intestines, etc.), as well as 25 **liver tissue** regions, which in turn are divided in 10 healthy areas (normal, vessels and border areas) and 15 tumor areas ( 6 fully inside, and 9 border areas ).

On a first phase, the *characterization*, the spectrums for some of those areas have been represented individually, and progressively they have been gathered together in order to compare the differential characteristics of the affected areas. Both multifractal an iso fractal spectrums were analyzed utilizing different approaches. For the sake of concision, only the most relevant experiments are going to be reviewed.

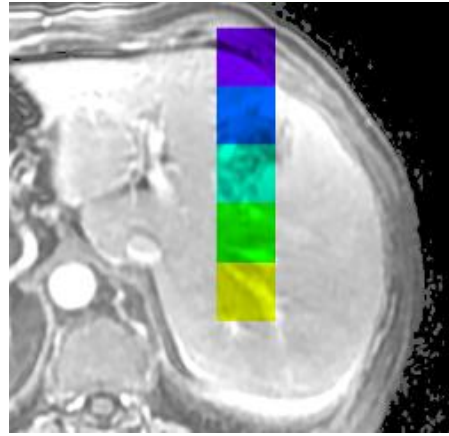
The second phase is the *descriptor definition*. Bearing in mind the characterization experiments, the most discriminating properties observed in the spectrums are converted in a set of scalar descriptors, which retain the classification information.

During the third phase each one of the descriptor was calculated for the 35 ROI set. The results were conveniently plotted, trying to find out which descriptors are able to separate the metastasic tissues from the healthy tissues. Finally, the best descriptors were combined, in order to obtain the most discriminatory tissue-classifier.

## 8.2 Characterization phase

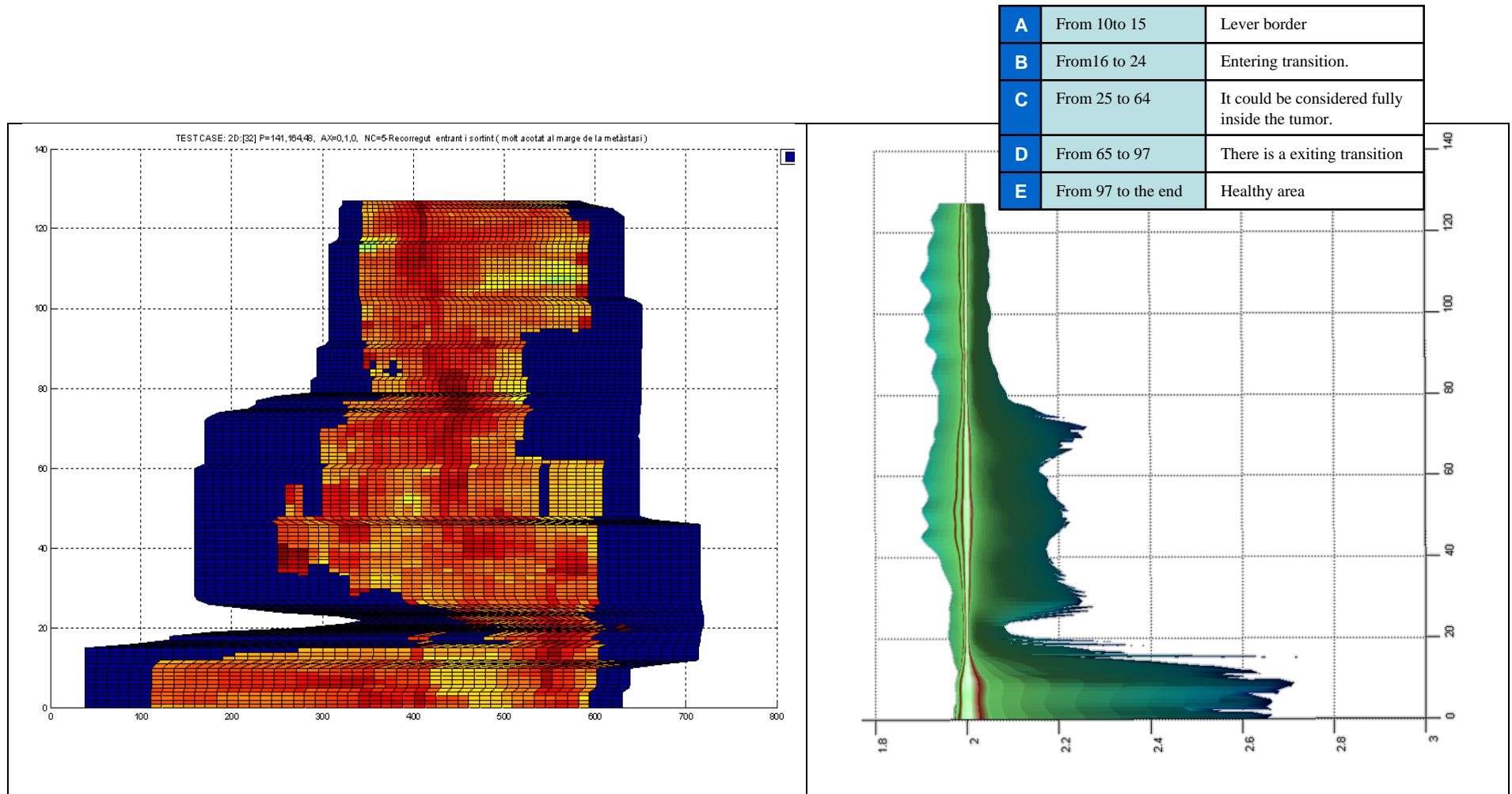
### 8.2.1 In-and-out scanlines

Before proceeding to the analysis of specific test case areas, one of the first contacts procedures has consisted of the spectrum evolution observation while crossing the affected areas. In these scan-lines, it was expected to detect some determinant differences that were easily translated into selection parameters. Figure 8.1 shows the sweeping path performed by a 32x32 sliding window. The track is made of 128 steps (4x32), entering and leaving the metastasis area. Only the first box (dark blue) and the queue boxes (green and yellow) are neatly outside the tumor. Figure 8.2 shows the multifractal and iso fractal spectrums through the path. There, it could be noticed that exists a clear relationship between the broadness and the density range existing in the regions.



**Figure 8.1:** Scanline crossing the metastasis- The path starts in the dark blue box, and ends in the yellow one.





**Figure 8.2: Isofractal (left) and Multifractal spectrum (right) of the scanline in Figure 7.1**

Path segments B and D, correspond to homogeneous liver tissues, with a small range of densities, and thus, smaller spectrum width.

Multifractal in tumor area (C) is wider; this is mainly provoked by a growth of the right slope. In contrast, the left slope seems to keep its width similar to the healthy areas, but the shape it is changed, since the position of  $\alpha(0)$ ,  $\alpha(1)$  and  $\alpha(2)$  (marked in red) start to diverge when entering the tumor.

In a similar manner, the isofractal spectrum grows wider in C. Its shape seems to be characterized by a single centered peak, but it is not applicable in the entire path. In this case it is quite difficult to establish a characterization rule, due the complexity and variability of the curves.



## 8.2.2 Multiple area comparison

### 8.2.2.1 Isofractal Spectrum

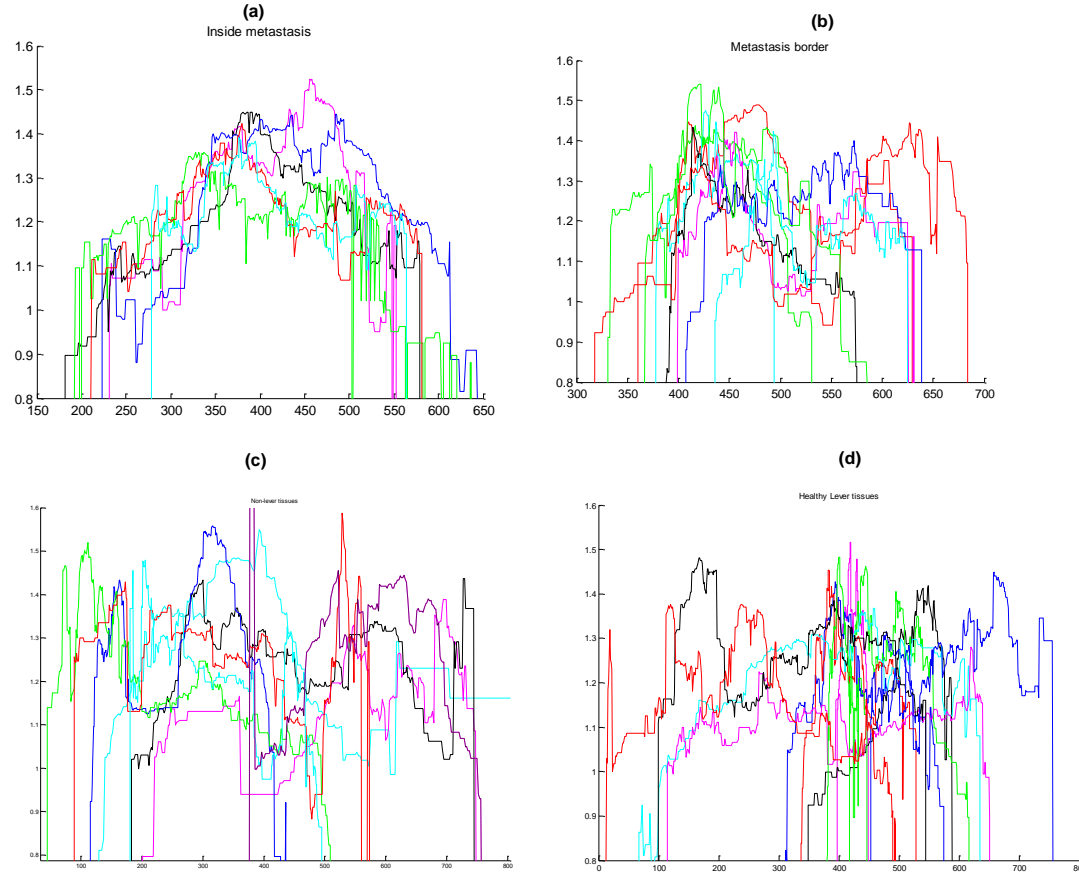


Figure 8.3: Isofractal spectrum of different ROI.

One of the limiting aspects of iso-fractal spectrum is that some of the information available is related to an absolute density value. Since the gray scales in MRI are “volatile”, being highly dependent of the scanning conditions, it is not convenient to use descriptors resulting in absolute density values when characterizing the spectrums.

Figure 8.3 is a collection of curves from affected areas (a), border areas (b), healthy liver (c) and other tissues (d). It has been quite difficult to find a pattern to characterize the tumoral areas. First and second order statistic operators have proved to be unable to provide acceptable classification; mean, variance and maximum are not able to differentiate the tumoral curves from the others. This could be seen in Figure 8.5, where discontinuous lines represent the tumor areas.

Fourier analysis has been also attempted without results.

One of the characteristics that seem more promising is the width of the spectrum, furthermore, the broadness of spectrum above a certain FD threshold. It seems that the tumoral areas have a bigger fragment of spectrum on the high range of fractal dimension. It could be interpreted as a sign of high irregularity at multiple levels. In any case, this characteristic is far from being sufficient for classification in the case of a single image.

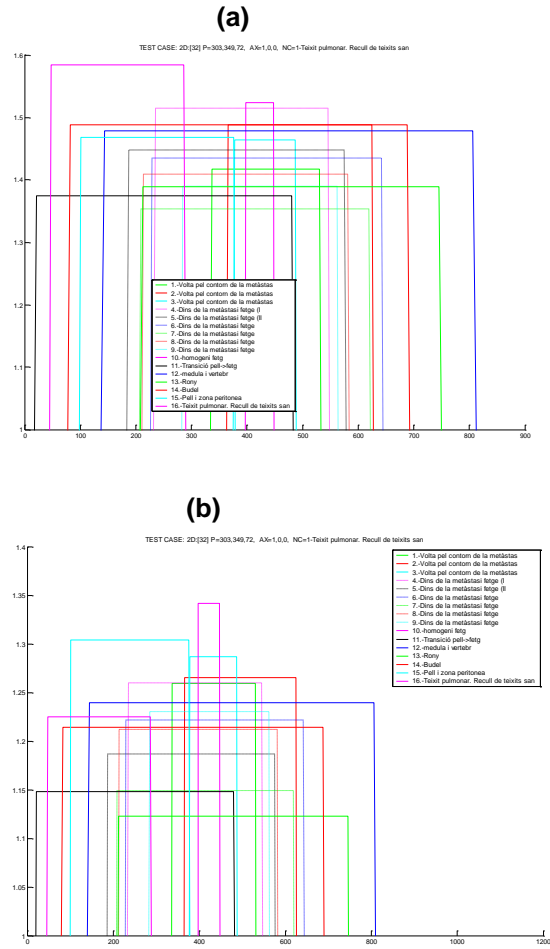


Figure 8.5: mean (a) and max (b) operators applied to ROI. Discontinuous lines indicate tumoral areas.

After some careful observation, it can be noticed that tumoral spectra is characterized by a triangular shape. Following that thread, the spectrum is preprocessed using a low pass filter kernel of size 3 in order to get rid of oscillations, and the program calculates the standard deviation of the error committed by the approximation of the spectrum with a pyramidal signal. The pyramid is situated on top of the value 1.0 of Fractal Dimension, its width is the 85% of the analyzed spectrum, and the peak placed equidistant from the extremes, having a height equivalent to the maximal Fractal dimension of the spectrum. The pyramid is slid and matched against the spectrum using the 15% marginal width, searching for the minimum error. Given such a definition, when combining the values of spectrum broadness above a DF threshold of 1.2, and the deviation error from the pyramid, we get a classifier that is able to successfully group the test case ROI, only two samples are erroneously categorized, a false positive and a false negative.

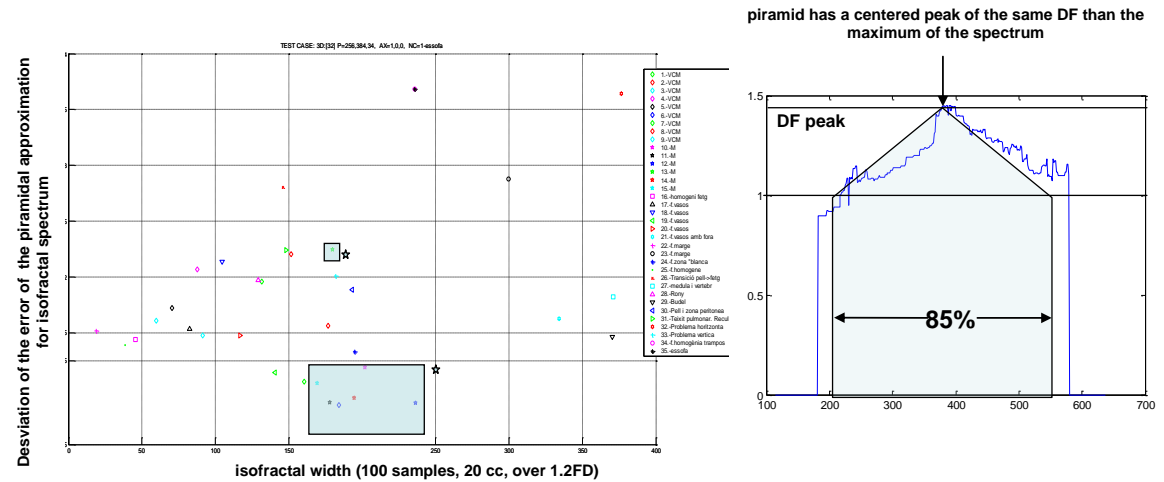


Figure 8.4: Deviation error of the approximation of the iso fractal spectrum using a pyramidal signal.

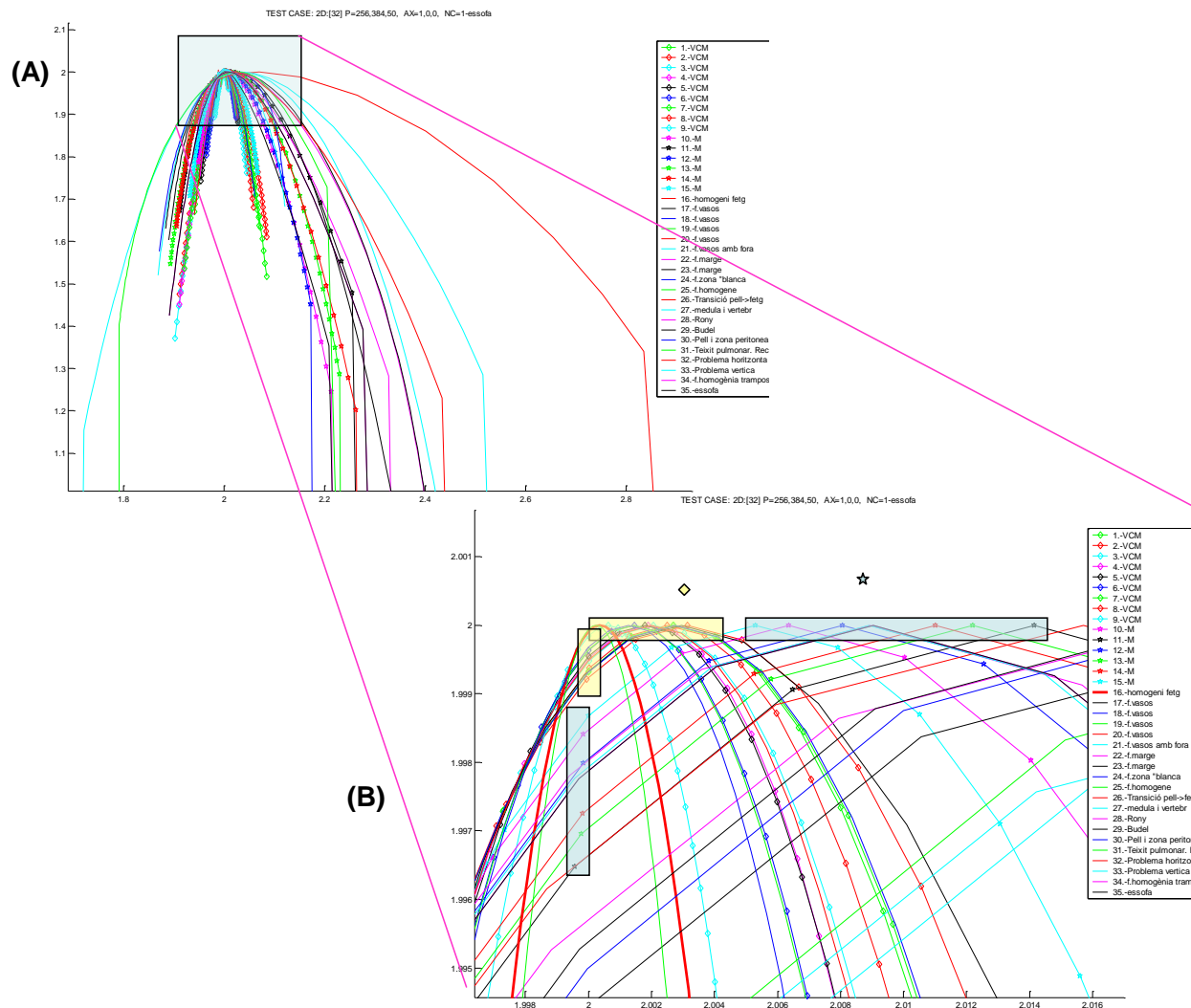


Figure 8.6: Mutifractal Analysis on 35 ROI

### 8.2.2.2 Multifractal Spectrum

Results obtained from Multifractal analysis of ROIs are quite resolutive. In Figure 8.6, the tumoral tissues have been marked with a star, and tumor border with a diamond. Normal healthy areas are straightlines.

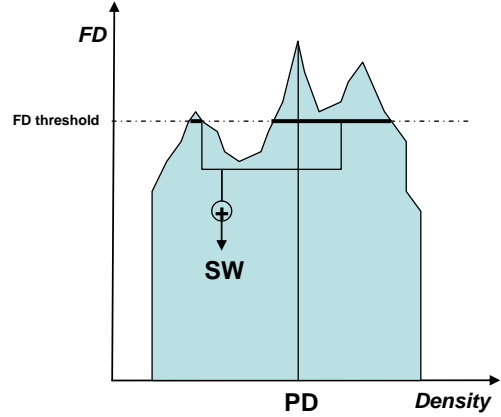
In the first overview image (A), it could be seen that the liver areas are conveniently grouped, which is a sign of a higher level of monofractality in liver when compared with other tissues. When focusing into the area occupied by liver curves (B), it could be noticed that there exists a separation of tumoral tissues from tumor border areas or normal areas. The healthy homogeneous areas (bold pink) show a great level of monofractality, observing that the enter process provokes a progressive displacement towards multifractality. The  $\alpha(0)$  becomes greater inside the tumor, as well as the displacement between  $f(0)$  and  $f(1)$ .

It also could be observed that the liver areas containing normal vessels fall in the same  $\alpha(0)$  range than metastatic tissues, which in fact is quite reasonable. In that case, the segregation could be obtained using the left slope of the spectra or the vertical balance.

## 8.3 Second phase: Descriptors definition

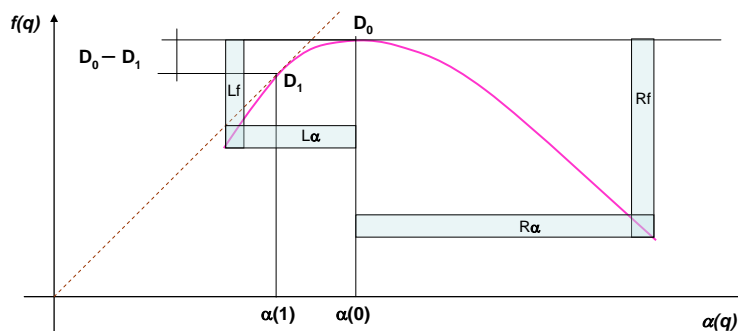
### 8.3.1 Isofractal spectrum descriptors

- Spectrum width (ISW).** The isofractal width contains information on the density range of a region. If we ask for the width over a certain fractal dimension threshold, it could be through as an indicator of global irregularity. The effect is enhanced when augmenting the value of the connected components filter.
- Density peak position (DP).** Surprisingly, the maximum FD is not significant for separating tumor areas. In the other hand, it may be handy when detecting pulmonary tissues, which have a remarkable high FD. In contrast, the peak density position seems to be meaningful for separating areas affected in the case of CT images.
- Pyramid Shape Factor (PSF).** As explained, the PSF is the result of matching the spectrum against a pyramidal-shaped signal. The pyramid signal is adjusted in function of the analyzed spectrum, adjusting its width to 85% of the spectrum, and the peak FD levered to coincide with the spectrum maxima. The standard deviation of the approximation error is calculated, obtaining PSF as final result.



### 8.3.2 Multifractal spectrum descriptors

- $\alpha(0)$  (A).** Marks the position of the spectrum maxima  $f(0)$ . It is quite sensible to variations of the overall image complexity.
- Spectrum Width (MSW).** Is the range of  $\alpha$  values which are occupied by spectrum samples. It is proportional to the variability of the image, and the range of different densities existing in the area.
- Left Slope (LS).** It is a measure of the fractality of the area, the greater the slope, the more monofractal behavior. Monofractal can be considered to be more regular or “pure” than multifractal, where multiple fractal sets are mixed.
- $f(0)-f(1) = (D0D1)$ .** It is a derived descriptor that has demonstrated to gather quite interesting information. In fact, it contains indirect information about spectrum and left slope, due to the basic properties of the fractal spectrum.
- Horizontal and vertical balance (HB / VB).** These parameters are used to represent the shape factor which permits capturing the general proportions of the spectrum with respect the apex  $\alpha(0)$ .



$$LS = \frac{Lf}{L\alpha}$$

$$HB = \frac{R\alpha}{L\alpha}$$

$$VB = \frac{Rf}{Lf}$$

**Figure 8.7:**  
Multifractal  
spectrum  
parameters

## 8.4 Third Phase: Classifier definition

### 8.4.1 Static classifiers

The static classifiers have demonstrated a low to medium performance on the arterial and retarded images. As will be seen, best results were achieved using multifractal descriptors.

#### 8.4.1.1 Isofractal static classifier

The quality of selection using the isofractal descriptors already detailed is not sufficient, and more work will be required in order to obtain better results. The best selection was reached using the following combination:

$$C0 = \begin{cases} IS1: PSF < 0.015 \\ IS2: 150 < SW < 250 \end{cases}$$

Isofractal spectrum generation is to be done using windows of 32x32. The size of the windows selected for fractal study is 32 voxels. This value has shown a good balance between spatial and fractal precision but, again, the size is highly dependent on the target area size. The algorithm has been adjusted to a resolution of 100 samples, and the threshold for the filter of connected components has been set to 20 voxels.

#### 8.4.1.2 Multifractal static classifier.

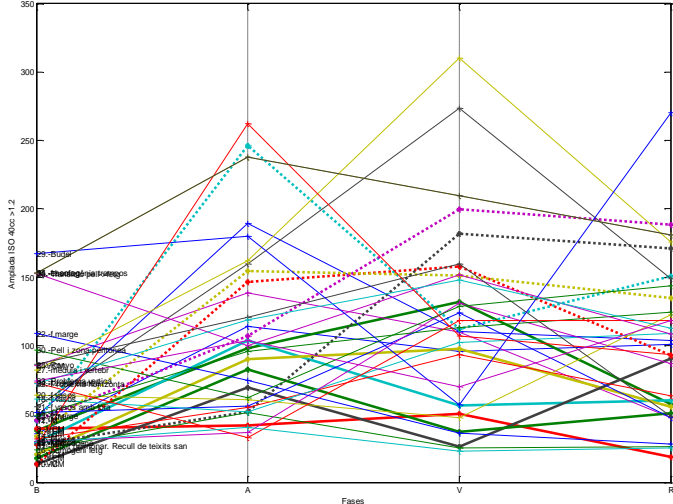
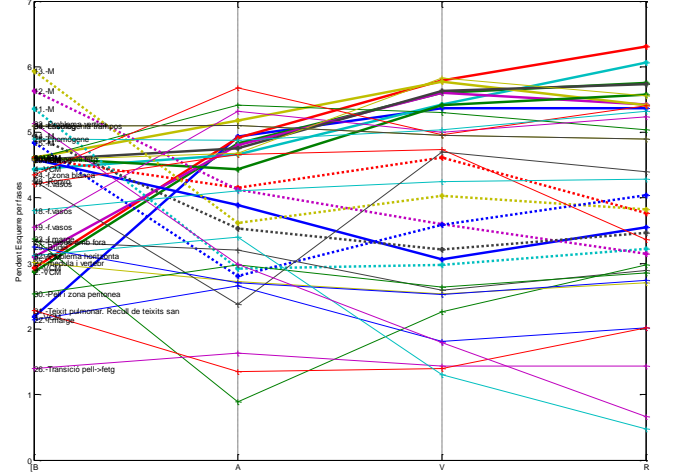
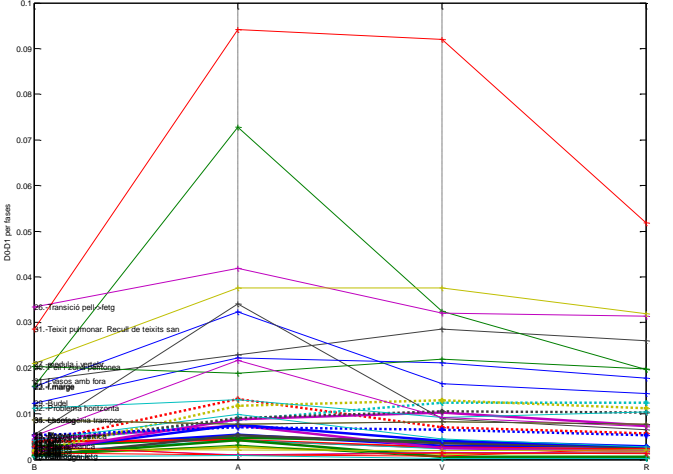
The characteristics of multifractal spectrum for retarded static MRI are remarkably captured using D0D1 and SW. The result is enhanced when constraining the HB and VB. The classifier could be defined as follows:

$$C1 = \begin{cases} MS1: 0.005 < D0D1 < 0.0135 \\ MS2: 0.13 < SW < 0.47 \\ MS3: HB < 4 \\ MS4: VB > 1.5 \end{cases}$$

The window size of the multifractal filter is set to 32x32, the same used in isofractal analysis. Both sizes are kept the same for equaling the capability of resolution in space and linear regression accuracy.

### 8.4.2 Dynamic classifiers

The analysis of dynamic parameters has been performed by observing the value of each parameter along the different image stages. In next table the time evolution of some of these parameters is shown. Each plot contains the value of the observed parameter for each of the 35 ROI. Discontinuous lines belong to marked tumoral zones, while bold lines are tumor borders.

	<p><b>ID1 ( based on ISW )</b></p> <p>The tumor areas show a differential behavior regarding ISW descriptor. Most of the ROI suffer an important gain from basal to arterial image, but most of them fall again for retarded image, with the exception of tumoral regions, which kept part of the gain or even increase its value even more.</p> <p>The marker works much better if the gain is defined as percentage:</p> $ID1 = (ISW.R - ISW.B) / ISW.B$ <p>For <math>ID1 &gt; 2.0</math>, the area is selected.</p>
	<p><b>MD2 ( based LS)</b></p> <p>Tumor areas show a clear descend on LS from basal to arterial image, which imply an increase of multifractality due to tumor vessel illumination.</p> $MD2 = (LS.A - LS.B)$ <p>For <math>MD2 &lt; -1</math>, the area is selected.</p>
	<p><b>MD3 (Based on D0-D1)</b></p> <p>D0-D1 shows a great raise-up for arterial phase in most of the tissues, and a progressive recovery of the original status, achieved at retarded phase. Tumoral tissues show a small but continuous growth which differ from other areas.</p> $MD3 = D0D1.R - D0D1.B$ <p>For <math>MD3 &gt; 0.003</math>, the area is selected.</p>

## 9 MRI CLASSIFIERS AND RESULTS

As a kind of conclusion, this chapter refers to the three classifiers that obtained the best results. The performance of the classifiers is 100% in the ROI classification studies, but it is far from perfect in the final volume visualization, where the entire volume is analyzed.

Although there is one of the classifier that excels, the others have been included as representatives of the usage of a restricted range of descriptors.

Additionally, the visualization part and the mechanism of tumor highlighting is described.

### 9.1 Final classifiers

A large number of classifiers may be defined by the combination the descriptors already introduced. For the sake of simplicity, it is interesting to reduce at maximum the number of descriptors required for the classifier. Therefore, a choice of 2 descriptors is quite convenient, but unfortunately it is not possible to fulfill that requirement in some scenarios.

As it has been already presented, descriptors may fall into static/dynamic and iso fractal/multifractal categories. Several classifiers have been built, using different combinations of classifier categories. In first place, static descriptors have been tested, trying multifractal and iso fractal independently and in cooperation mode. Later on, dynamic descriptors were tested. Finally dynamic and static were combined.

From the results of the tests, it has been verified that the most successful classifiers are:

- $C1 = \{ MS1, MS2, MS3, MS4 \}$  : Just static monofractal parameters.
- $C2 = \{ MS1, ID1 \}$  : One static multifractal and a dynamic iso fractal
- $C3 = \{ DS1, DD2 \}$  : Both dynamical multifractals.

All the classifiers are successful in the separation of the tumoral areas in the 35 ROI set used in the characterization phase. Figure 9.1 and Figure 9.2 show the results of applying  $C2$  and  $C3$  over the ROI testing set. The metastatic tissues, marked with a star, could be found grouped in the blue boxes.

In spite of the positive results, once the classifiers are applied to the whole volume, some healthy areas appear highlighted as tumoral. Hence, the final quality criterion was set to the incidence of false positives and the sensibility to surrounding tumoral areas in the final volume visualization. More details on this phase are covered in the following section.

$C2$  is, by far, the best classifier. It provides the smaller number of false hits, and it is really accurate in the description of the tumor area while using a reduced number of parameters. The counterpart is that requires multiple images from the contrast MRI studio.

In contrast, the main advantage of  $C1$  is that works in a single image, though it requires a high number of descriptors, which make it difficult to parameterize in the hypothetic case of application in different images. Furthermore, it has multiple false positives on the liver borders and in external regions.

$C3$  restricts the filtering to the usage of multifractals, but it seems that there is not information enough to provide such discriminative results as in  $C2$ .

## 9.2 Region voting and visualization

The selection process works by applying a classifier on a sliding window inside the plane of a slice. If the result is positive, all the voxels within the region are given a vote. In order to make use of spatial coherence, the adjacent voxels lying in the previous and next neighboring slices are also voted. Each voxel may receive multiple votes, coming from overlapping sliding windows placed in different positions on the same slice, or from adjacent slices.

In the visualization phase, the volumetric model is shown by using the default transfer function as usual. Additionally, the voxels which exceed a certain vote threshold are highlighted. The visualization is quite simple, the higher the number of votes, the higher the selection index. This index is transformed using a color map in a hue, which is applied to the voxel, respecting its luminance value.

The application let the user to specify the minimal number of votes for considering a region selected, which permits adjusting the sensibility of the results.

The results could be displayed in 2D or 3D, though tridimensional visualization requires more adjustment effort. In Figure 9.3 the classification results for C1, C2 and C3 are shown in 2D, using a selection color is in the range of reds. As has been advanced, the C2 performs better than the other classifiers.

In the following figures, 9.4, 9.6 and 9.5, the results of C2 are shown using volumetric visualization. In this case, the selection color is in the range of blues. In those figures, the “*tumoral cave*” structure within the liver could be noticed.

## 9.3 MRI Classification Results

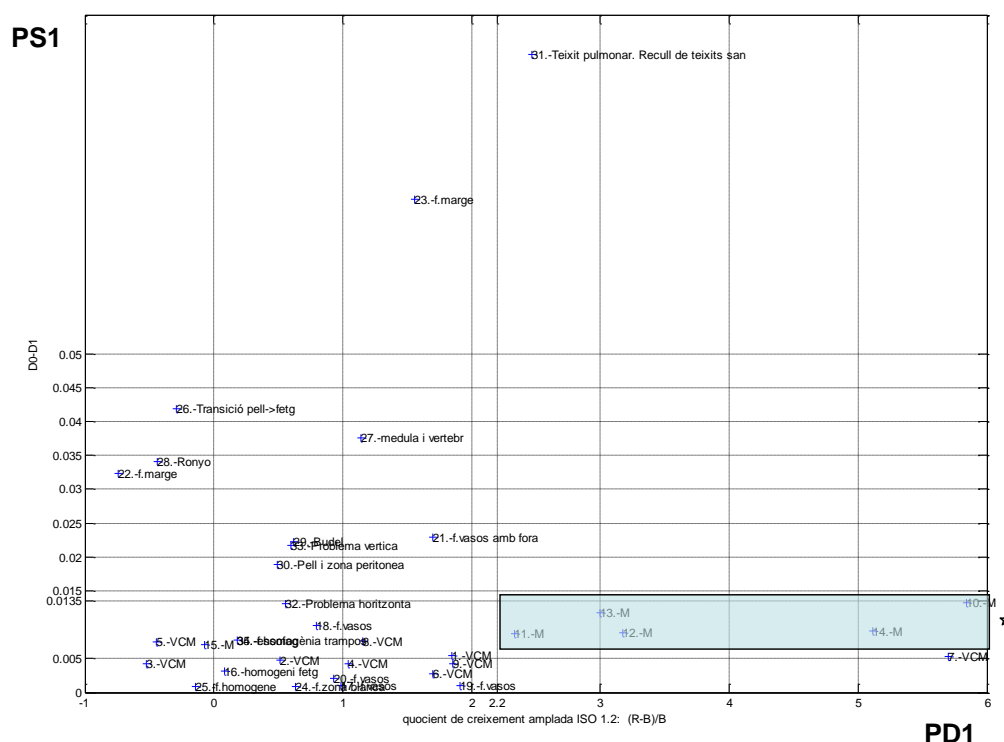
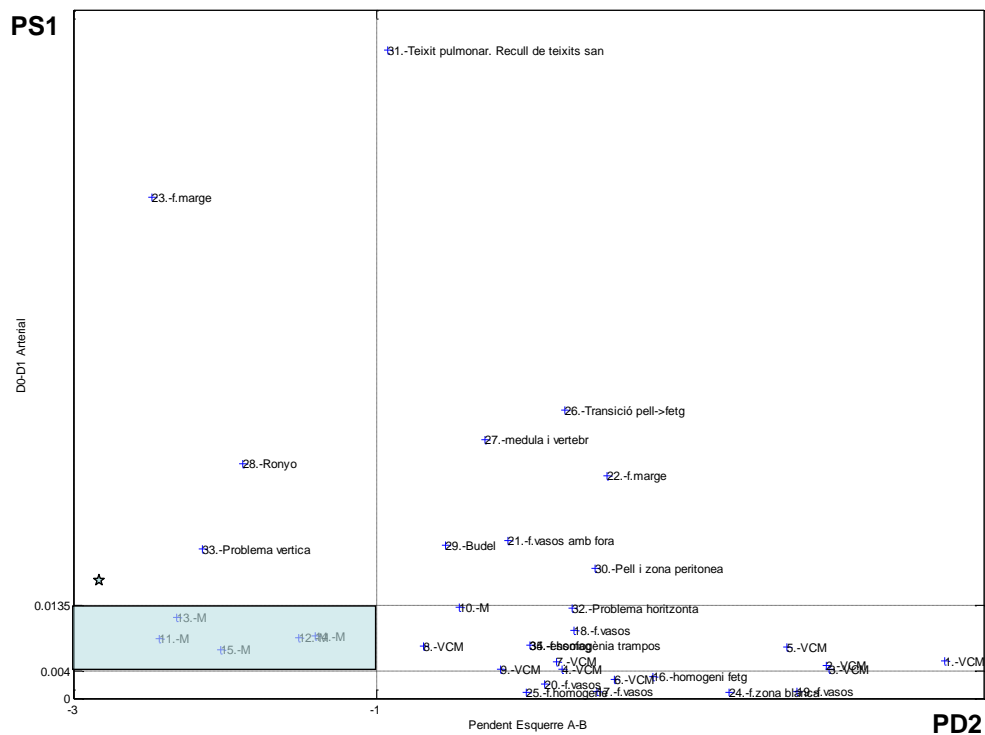


Figure 9.1: Classifier C2: DS1 versus DD1. Tumoral areas marked with letter M fall grouped in the blue box.





**Figure 9.2: Classifier C3: DS1 versus DD2. Tumoral areas marked with letter M fall grouped in the blue box.**

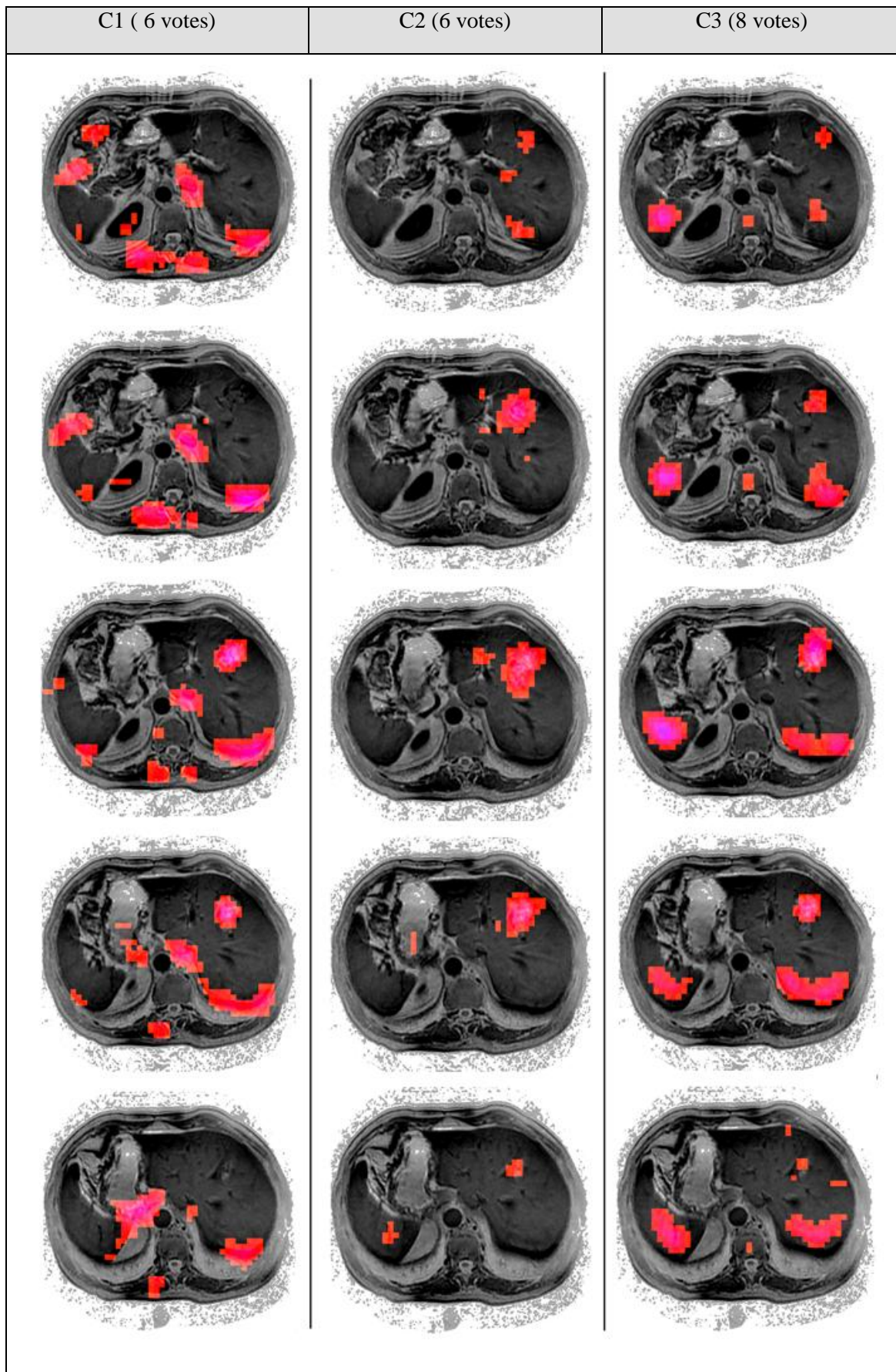
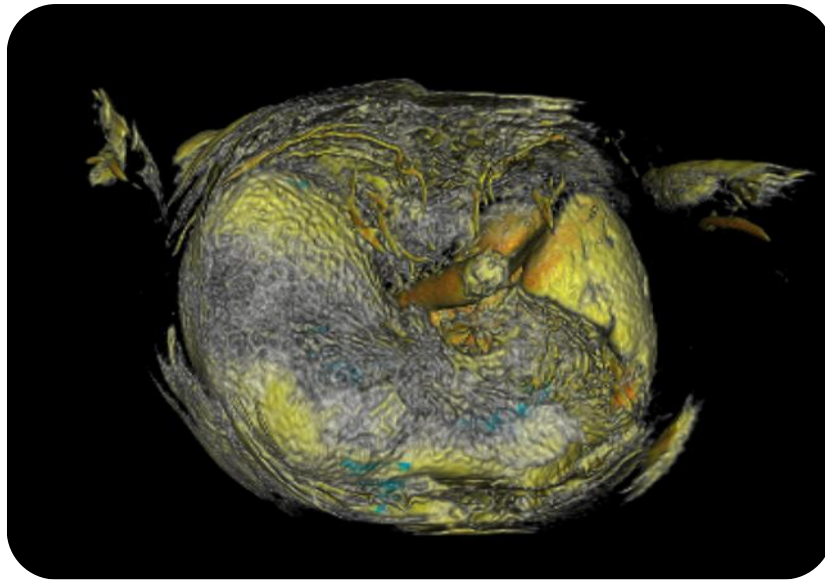


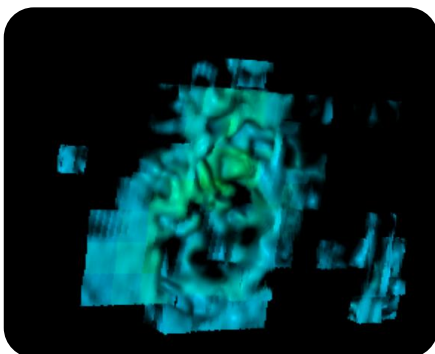
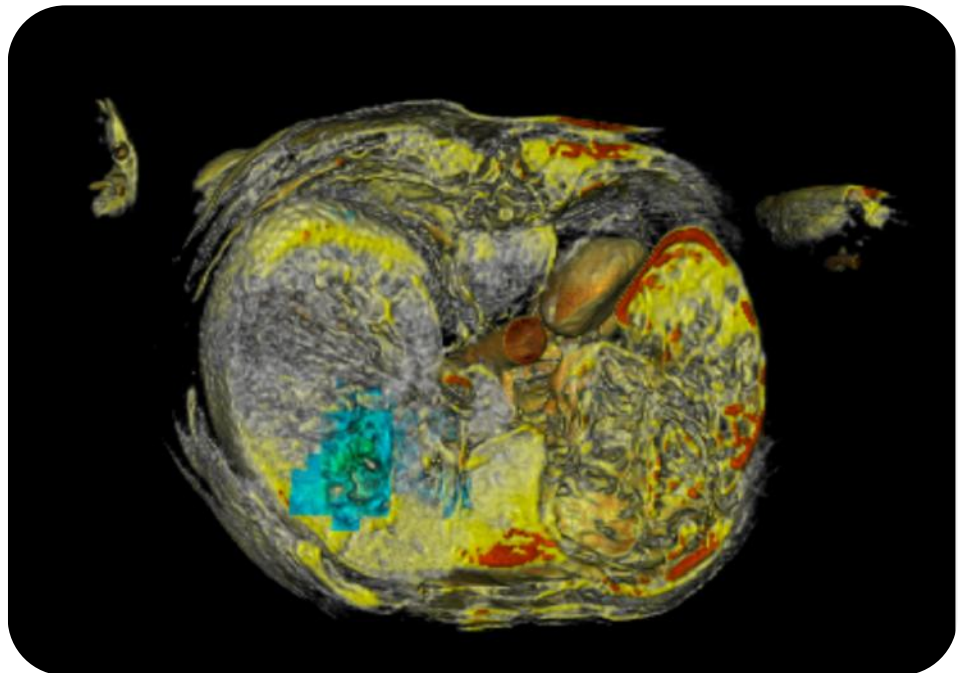
Figure 9.3: Classifiers applied to MRI test set A. MRI images has been inverted for an improved visualization.



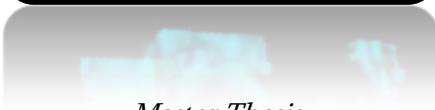
**9.4: Volumetric representation of Liver area. Blue areas are those marked as tumoral by C2 classifier.**



**9.6: Cross section of the previous volume, showing clearly the blue metastatic tissue selected by the C2 classifier.**



**9.5: The image has been obtained hiding all the healthy tissue. It could be appreciated the irregular cavity of the tumor.**





# 10 ANALYSIS AND CLASSIFICATION OF CT BASED ON FRACTAL CHARACTERIZATION

## 10.1 Introduction

Despite the thesis core is devoted to MRI analysis, this chapter intends to summarize some of the works done with CT images following a similar methodology. It is interesting to realize the notable differences between CT and MR imaging, and to know how the already presented techniques behave with CT images.

For such purposes, the dataset B has been explored. This image can be considered a difficult challenge for a radiologist. The metastasis is in a preliminary stage, it is relatively small, and the image is a routine control CT and, hence, low-resolution. Given the metastasis shape, it is barely visible in a **unique** slice (#135), and could be easily missed.



Figure 10.1: CT #135 slice showing a metastasis area circled in red.

In fact, the visualization problem here it is due to the fact that the increment of the density of the metastasis tissue is so small when compared to the whole dynamic range that is lost in terms of visual perception. The solution here is quite straightforward if applying a drastic restriction of the dynamic range, focusing to the levels of interest. The result, after applying a color mapping is shown in Figure 10.2.

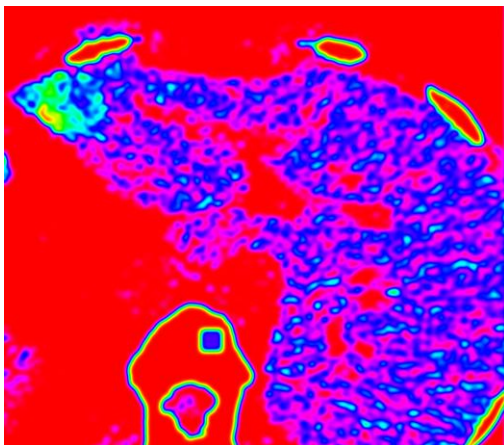


Figure 10.2: CT slice #135 after applying a range limitation and a color map.

In the processed image, the affected area is clearly highlighted. If the same process is applied to the neighbor slices, it reveals some nearby areas that may be fragments of the same metastasis that were not visible by the naked eye. However, the results in those slices are not relevant by themselves, only are remarkable when taking into account the vicinity to the reference slice. The results are shown in Figure 10.3.

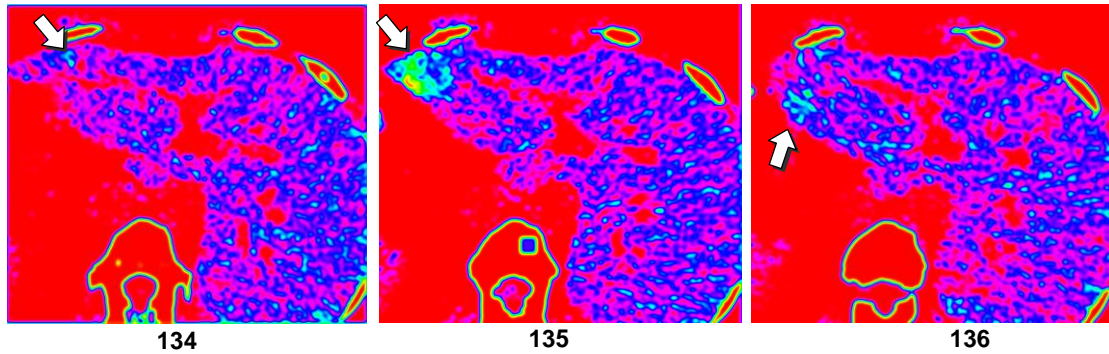


Figure 10.3: Metastasis is clearly visible in the slice 135. Expanding the dynamic range also enables the detection of traces in the surrounding slices (134 / 136)

### 10.1.1 Fractal analysis characterization

#### 10.1.1.1 Multifractal analysis

From the point of view of fractal analysis, the techniques already presented have shown little applicability. The multifractal spectrum is not able to separate affected areas from normal homogeneous tissue. It happens that the texture differences from metastatic tissue to normal tissue are not perceptible from the multifractal approach. In the Figure 10.4 could be appreciated that the spectrum corresponding to tumoral area (green-starred) is not separable from other homogeneous healthy areas of liver (series 3,4 and 5). The spectrum main parameters such as width, slopes and  $\alpha(0)$  are identical. The unique differentiating parameter is the spectrum width, but as could be seen in the left figure, it is far from being selective enough.

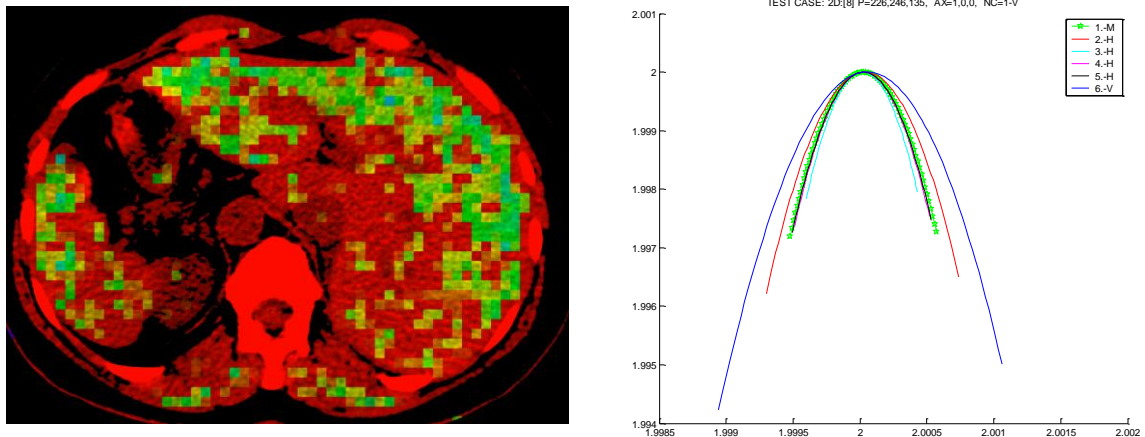
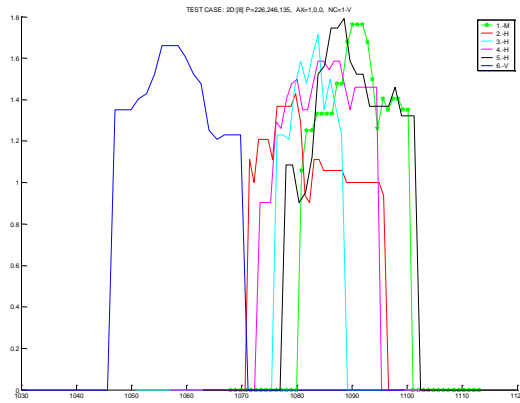


Figure 10.4: (Right) iso fractal spectra for several areas inside the liver. The green starred line represents the metastatic tissue. It could be appreciated that is quite indistinguishable from other areas. (Left) The width of the spectrum has been used as spectrum descriptor, with non-acceptable results.

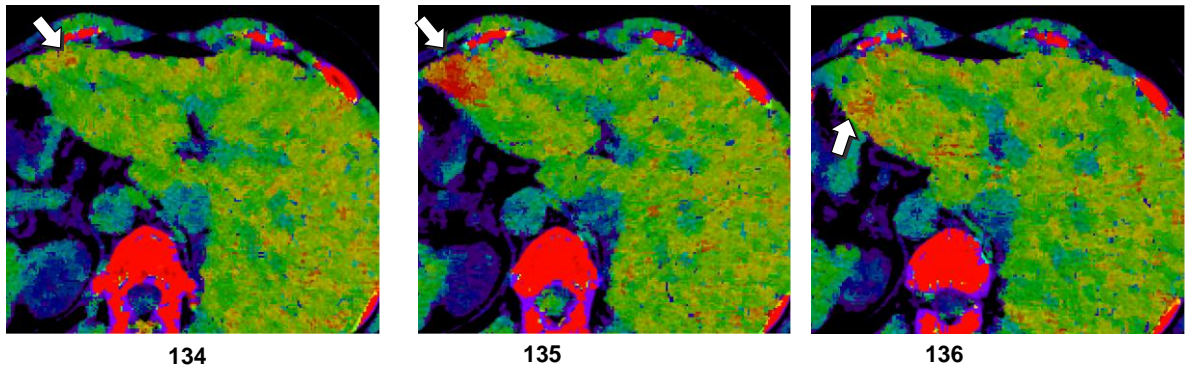
### 10.1.1.2 Isofractal analysis



**Figure 10.5: Isofractal spectra of CT liver areas. Green starred is inside the tumoral area, while the others are healthy zones.**

In that case, isofractal spectrum is able to obtain better results than multifractal spectrum since the most relevant information is in the underlying density level. It could be seen in Figure 10.5 that the spectrums for the areas inside the metastasis (green starred) are found displaced to the high density band. Besides that, there are no remarkable features in the shape of the spectrum, neither the width nor the height. The first and second order statistics do not give further information. The unique property that provides some classification potential is the peak position, which corresponds to the density value with a higher fractal dimension. Using that descriptor we

obtain a classifier with similar results to those given by the dynamic range compression. Anyway, the fractal approach seems to be slightly more sensible and accurate than the direct range manipulation as shown in Figure 10.5.



**Figure 10.6: CT slices showing the density value corresponding to the peak of the spectrum. Window size used is 8, and the connected component filter is set to 10.**

At this point, we may ask if the extra computational cost required by isofractal technique is worth the effort for this slight visual gain. The answer would be probably negative in most of the cases.

## 11 CONCLUSIONS AND FUTURE WORK

In this work the construction of a liver metastasis detection system on MRI datasets has been presented. The detection algorithm relies on Fractal Analysis for obtaining the characteristics of the tissues. In previous works it has been demonstrated that Fractal Analysis is a powerful tool for cancer detection and staging, and multiple techniques and variations have been presented. In this work, two different approaches have been taken, one of them based on the well-known Multifractal Spectrum, and a new contribution, the Isofractal Spectrum. Both techniques have demonstrated to be in some way complementary.

Isofractal spectrum has been run against a testing set in order to determine its properties and weaknesses. Given its geometric nature, the technique is sensible to the presence of high complexity isocurves within the volume, but its precision is affected by the lack of spatial resolution, which limits the reliability of the fractal dimension calculation through box counting. In order to soften the negative effects of additive curves in the regression slope calculation, it has been proposed the usage of a component connected filter to reduce the noise contribution.

A group of selected regions of interest within an MR image have been studied, by comparing the shapes and features of the fractal spectrum for malignant and benignant tissues. This survey permitted the definition of two set of descriptors, one for each of the fractal techniques. Multifractal descriptors were already known, though this work adds two new ones, horizontal and vertical balance, which has been useful in the static characterization of the signal. Isofractal descriptors are entirely original.

The phase of spectrum characterization was initially performed in a static image. In that sense, multifractal descriptors have demonstrated better capabilities for tumor detection than isofractal. In a second phase, the temporal evolution of the MR contrast image was analyzed. The values for the descriptors were tracked in the time-line, studying the changes provoked by the contrast blood perfusion. Here, both techniques show good selection capabilities when comparing the basal phase values with arterial or retarded phase values.

Based on the characterization phase, the descriptors were combined to form classifiers. Those classifiers had been tested to successfully separate the tumor regions from the healthy regions in the 35 ROI testing set. When the detection scheme is applied in combination with the visualization pipeline, the system is capable of highlighting those areas classified as malignant. The voxels within the areas classified as malignant receive a vote. The number of votes is increased if the voxel belongs to several affected windows, which is a way to exploit the spatial coherence of the affected areas. Finally, the number of votes acts as the highlight weight factor in the resulting image.

Based on the visualization results, the final classifiers allowed the detection of the metastatic tissue, in spite of some false positive areas. The most proficient classifier has been proven to be the one using a combination of multifractal static and isofractal dynamic descriptors.

It is important to remark that the classifiers are applied to the entire volume, there is not a prior liver segmentation to reduce the working set. Hence, it is clear that the algorithm is quite sensible to certain properties only existing in the liver tumor.

The results points towards the suitability of multifractal analysis as a tissue classification tool, especially when the areas could be characterized by its textural irregularity properties. Regarding Isofractal Spectrum, it has demonstrated its effectiveness when comparing differential properties of an area in its time-line. When combining the two techniques, the classifiers benefits from both worlds, obtaining remarkable segmentation results.

On the other side, this work lefts multiple open issues, deserving a special mention the 3D analysis. Most of the work has been done in 2D. It seems reasonable that making the jump from 2D to 3D may provide a richer analysis and accuracy in the results, but the current lack of resolution make this approach quite difficult. The classifiers should also be tested on additional MRI datasets in order to generalize the results.

Another important point which requires further work is the characterization of iso fractal spectrum, probably using some other statistical analysis tool, or even fractal analysis, since the spectrum reflect some level of autosimilarity.

In relation with dynamical analysis, the amount of information that could be obtained in such kind of scanners is huge. Probably another type of image or volume analysis based of physic models of perfusion may contribute in a much better exploration of the images. In the other hand, some problems are derived from the time lapsed scans, one derived from the micro displacements of the person between MRI snapshots due to the respiration or other involuntary movements that may introduce distortion in time series. The other problem, which exclusively affects the MR imaging is due to the differences in the gray ranges between images for equivalent tissue densities. In such a case, it would be interesting to have some “prove areas” that permit equalizing the images corresponding to the same temporal series.

The master thesis has focused in classification using contrast MRI, but for cost reasons CT images are more commonly used. Liver CT captions are noisier and more difficult to interpret than MRI, which makes it a target of computer aided imaging. Further developments need to be taken in that direction.

Continuing with medical image types, it is usual that CT or MRI datasets also attach PET images. It would be of high interest that the systems provide a registered representation of information coming from different sources, or even more, the ability of displaying an historic evolution of the images of the same patient. Next generation aided-image analysis systems will need to take into consideration all these information to obtain reliable results.

More in the line of algorithmics, another interesting approach that could be explored is the usage of wavelet transformations for obtaining multifractal spectrum. The usage of DWT could probably improve the processing time and it could extend the power of multifractals by introducing the concept of spectrum directionality.

In the same direction, there are other fractal techniques to be tested, such as those related with Lacunarity, which basically deals with the study of the gaps between fractal sets. Fractal analysis could also be combined with other methods of image processing and pattern recognition, which may boost the accuracy of the results.




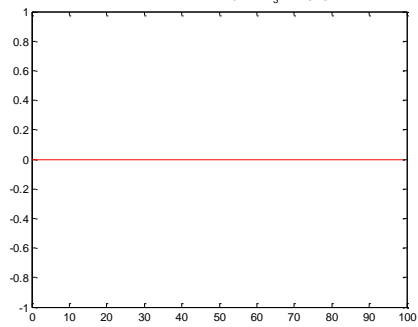
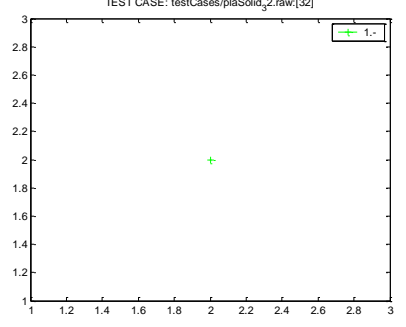

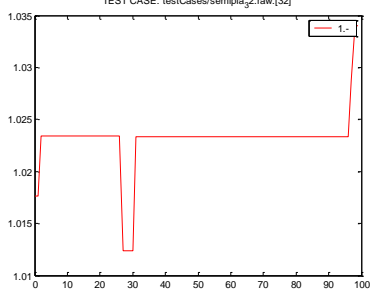
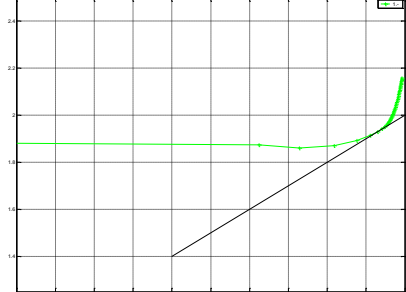

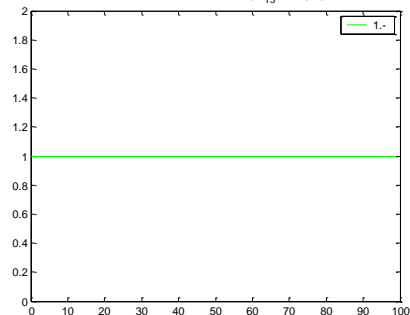
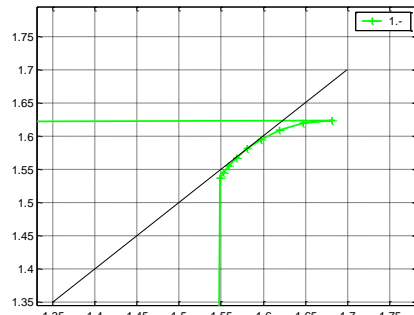
## 12 BIBLIOGRAPHY

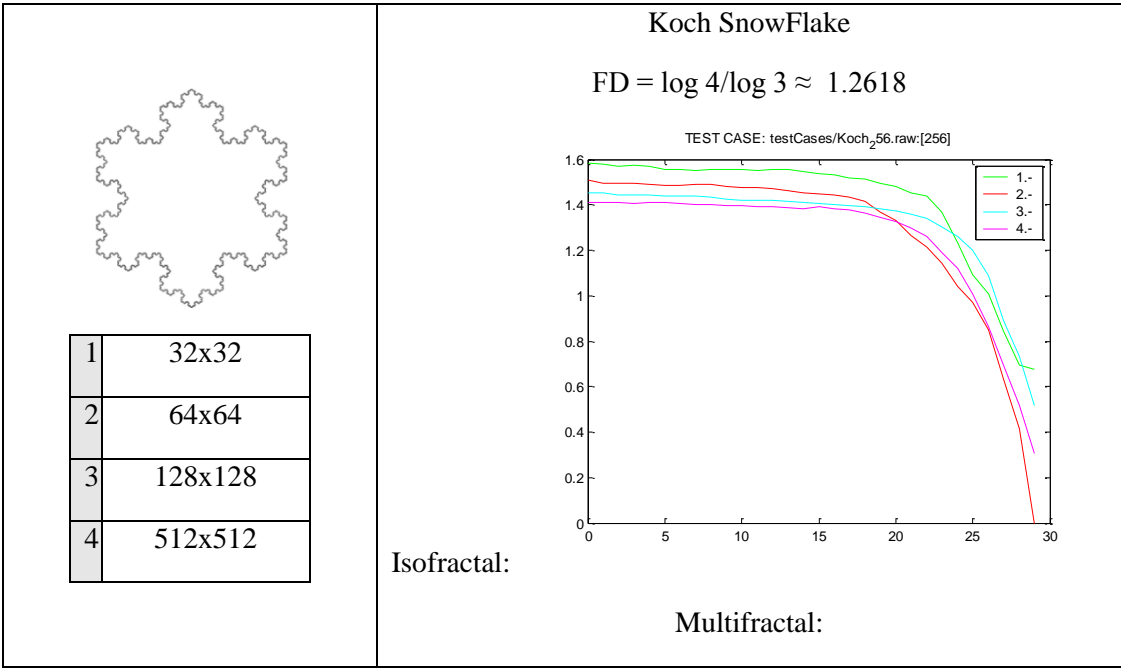
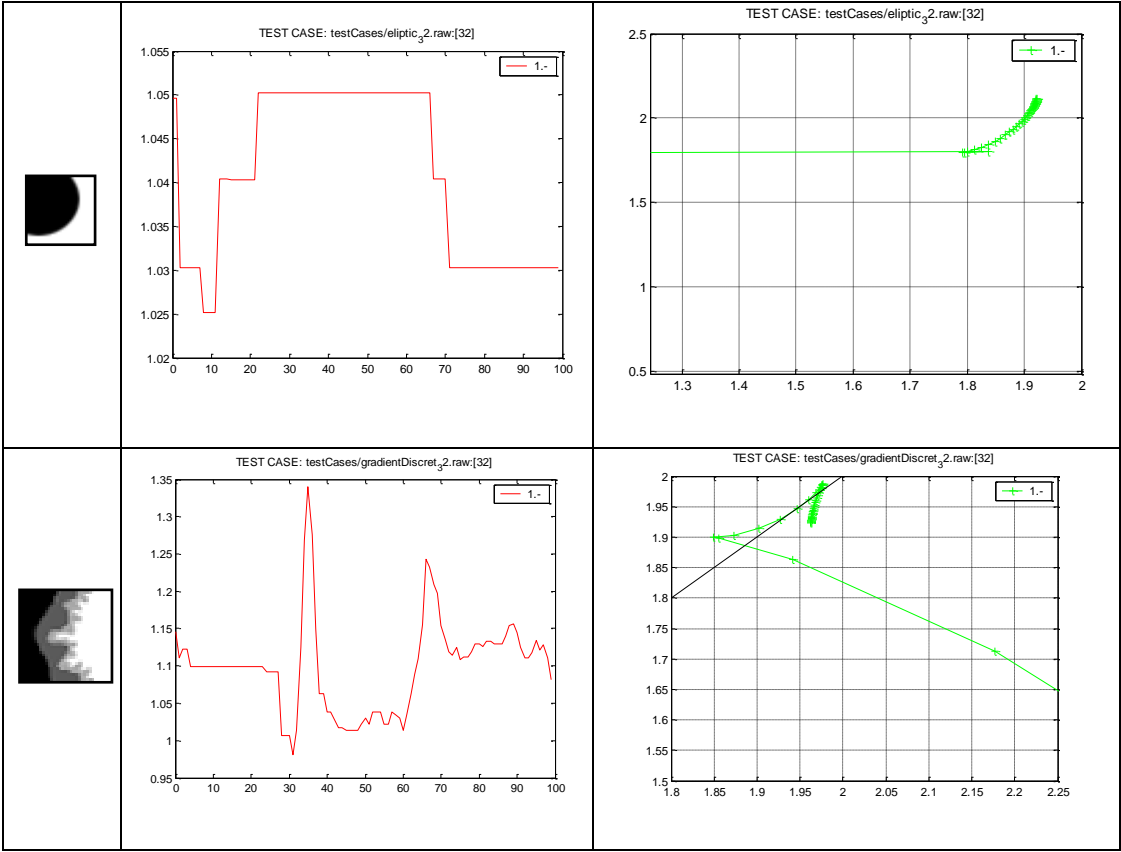
- [1] KLEIN J., BARTZ D., FRIMAN O. ET AL.: “Advanced Algorithms in Medical Computer Graphics”. *EUROGRAPHICS 2008*
- [2] AGUS M., GOBBETTI E., PINTORE G., ZANETTI G., ZORCOLO A.: “Real-time Cataract Surgery Simulation for Training”. *Eurographics Italian Chapter Conference* (2006)
- [3] CHO Z., JONES J. AND SINGH M.: “Foundations of Medical Imaging”. Wiley, New York, NY, 1993
- [4] HORNAK J.P.: “The Basics of MRI”. *Center For Imaging Science*. [web site] <http://www.cis.rit.edu/htbooks/mri/inside.htm>
- [5] MANDLEBROT B.B.: “The fractal geometry of nature”. W. H Freeman and Co. New York, 1983
- [6] KAANDORP J.A.: “Fractal modelling: Growth and form in biology”. Springer-Verlag, Heidelberg; 1994.
- [7] PEITGEN H, JURGENS H, SAUPE D.: “Chaos and Fractals”. Berlin: Springer-Verlag; 1992.
- [8] MICHAEL FRAME, BENOIT MANDELBROT, AND NIAL NEGER : “Fractal Geometry”. Yale University [web site] <http://classes.yale.edu/Fractals/Welcome.html>
- [9] A.CHHABRA AND RODERICK. V.JENSEN.: “Direct determination of the  $f(\alpha)$  singularity spectrum”. *Physical Review Letters*, 62(12):1327-1330, March 1989.
- [10] ARNÉODO A., DECOSTER N., ROUX S.G.: “A wavelet-based method for multifractal image analysis. I. Methodology and test applications on isotropic and anisotropic random rough surfaces”. *The European Physical Journal B*. 15, 567-600 (Springer-Verlag 2000)
- [11] DONGJIAO LV, XUEMEI GUO, XIAOYING WANG, JUE ZHANG AND JING FANG: “Computerized Characterization of Prostate Cancer by Fractal Analysis in MR Images”. *Journal Of Magnetic Resonance Imaging* 30:161–168 (2009)
- [12] AL-KADI. O.S.: “Texture Analysis of Aggressive and Nonaggressive Lung Tumor CE CT Images”. *IEEE Transactions On Biomedical Engineering*, Vol. 55, No. 7, July 2008
- [13] N. SARKAR AND B. B. CHAUDHURI: “An efficient differential box-counting approach to compute fractal dimension of image,” *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 1, pp. 115–120, Jan. 1994.
- [14] SONG-TAO LIU: “An improved differential box-counting approach to compute fractal dimension of gray-level image”. *IEEE 2008 International Symposium on Information Science and Engineering*.
- [15] BAISH J.W. AND JAIN R.K.: “Fractals and Cancer”. *Perspectives in Cancer Research. Cancer Research*, 60, 3683–3688, July 15, 2000.
- [16] FELDMANN A., GILBERT A.C., WILLINGERT W.: “Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic”. *Proceedings of the ACM/SIGCOMM'98*, pages 25--38, Vancouver, B.C., 1998.

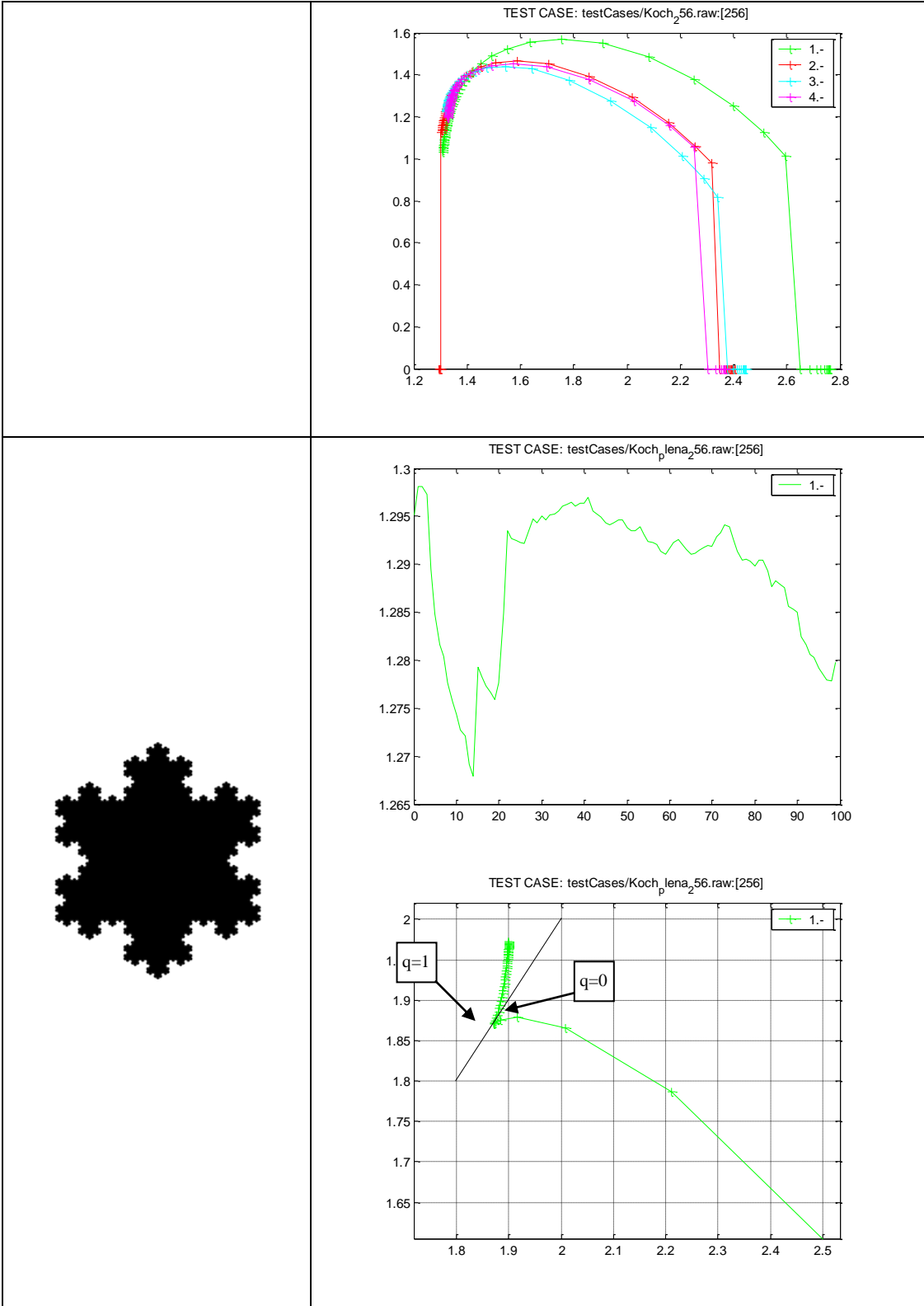
- [17] A.POSADAS, D.GIMENEZ ET AL.: "Multifractal Characterization of Soil Particle-Size distributions". *Soil Science Society of America Journal*. 65:1361-1367 (2001)
- [18] RELJIN IRINI S., RELJIN BRANIMIR D.: "Fractal geometry and multifractals in analyzing and processing medical data and images". *Archive of oncology* 2002 Volume 10, Issue 4, Pages: 283-293
- [19] RELJ I., RELJIN B., PAVLOVIC I.: "Multifractal Analysis of Gray-Scale Images". *Electrotechnical Conference, 2000. MELECON 2000. The 10th IEEE Mediterranean*. Volume 2, 2000 Page(s): 490 - 493 vol.2
- [20] STOJICA T., RELJIN B., RELJI B.: "Adaptation of multifractal analysis to segmentation of microcalcifications in digital mammograms". *Physica A: Statistical Mechanics and its Applications*. Volume 367, 15 July 2006, Pages 494-508
- [21] DERADO G., LEE K. ET AL.: "Wavelet-based 3-D Multifractal Spectrum with Applications in Breast MRI Images". *Lecture Notes in Computer Science. Bioinformatics Research and Applications*. Springer Berlin. 30/04/2008
- [22] RIEDI R.H.: "Introduction to Multifractals" *Technical Report*. Rice University. 1999

# 13 APPENDICES

## 13.1 Synthetic Test cases for multifractal and isofractal spectrum.

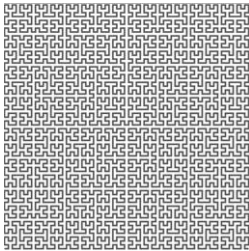
Shape (32x32)	Isofractal	Multifractal
		
		
		



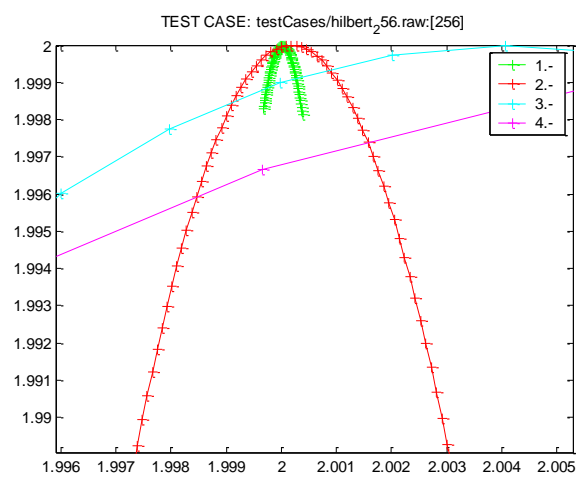
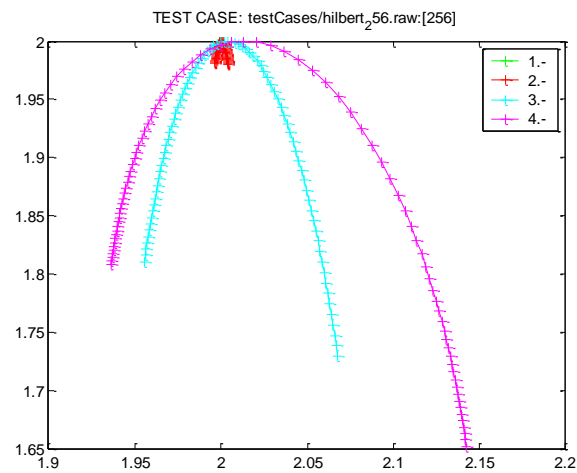
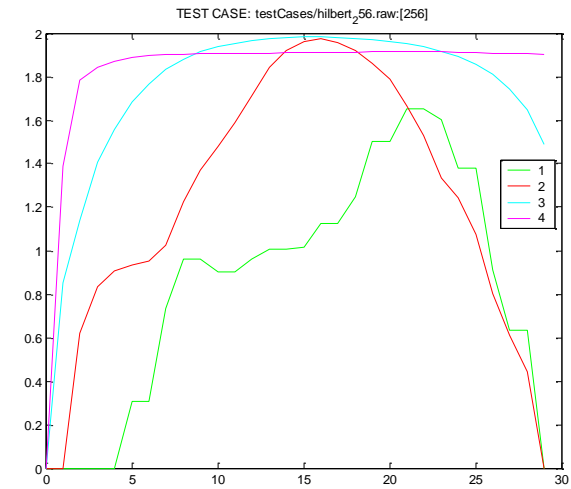


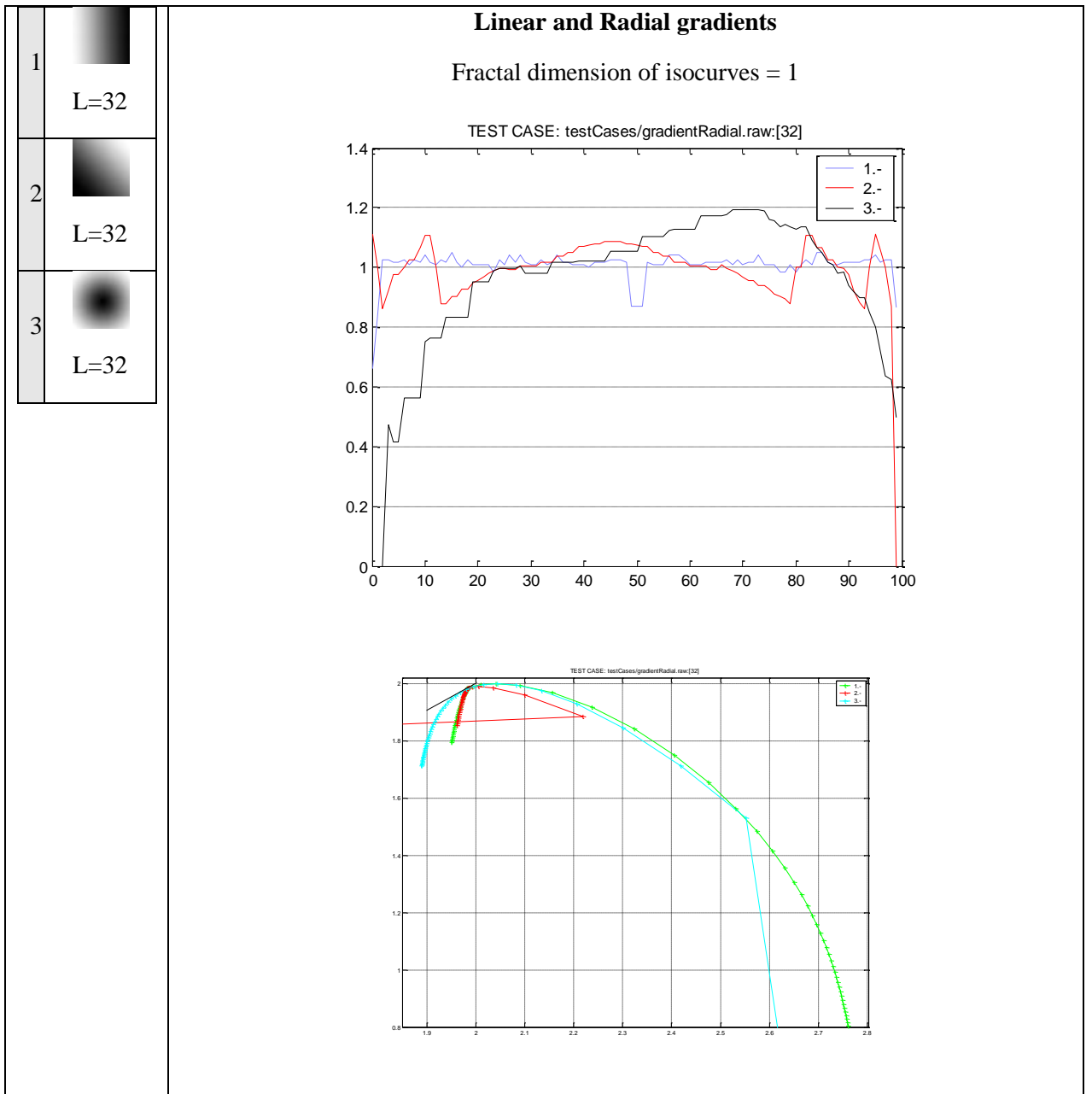
## Hilbert Space-Filling Curve

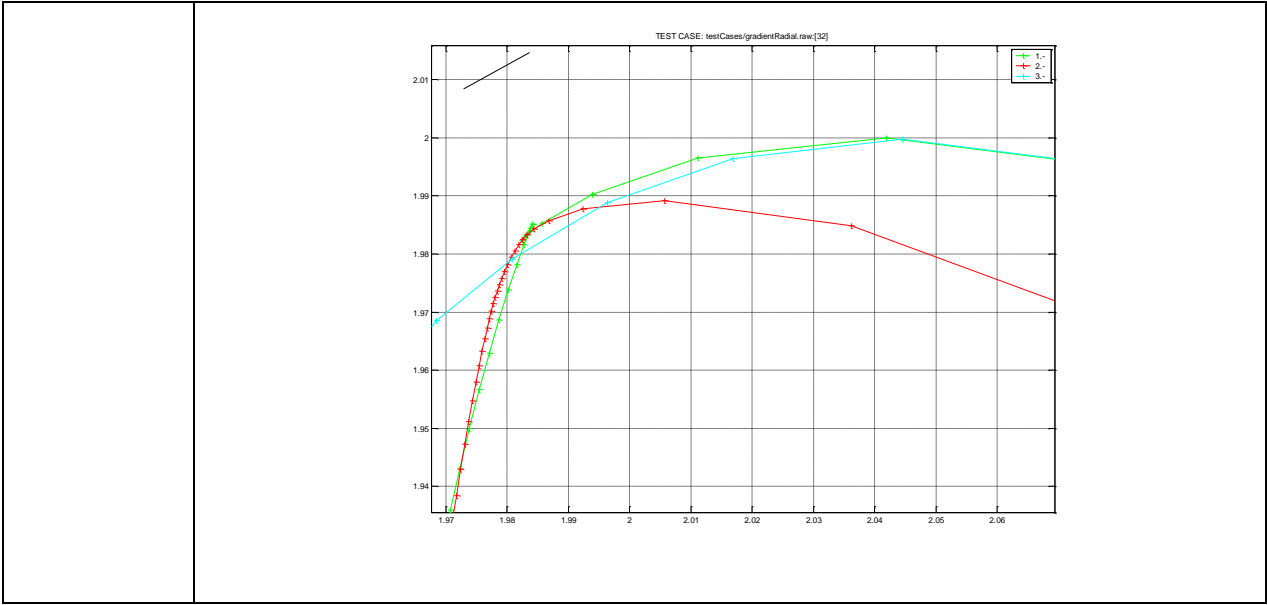
Fractal Dimension = 2



1	32x32
2	64x64
3	128x128
4	512x512



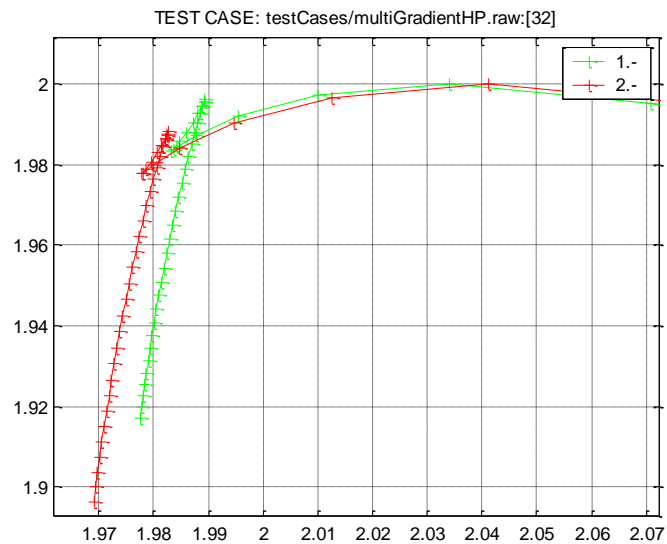
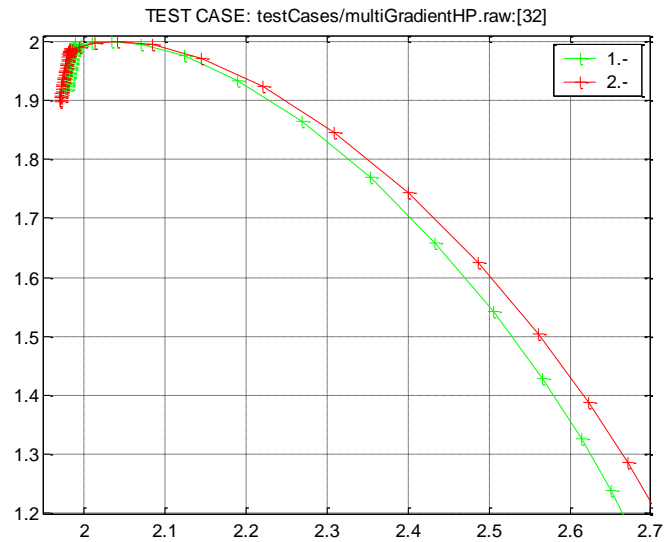




1		<p><b>Effect of low-frequency gradient on the image.</b></p> <p>Figure 1 shows a synthetic image simulating a branching tissue area. It could be observed that there is a low frequency grey gradient running on the image background, which alters slightly the perception of the structure. Despite of that, the tree structure could be still appreciated. In figure 2 it is shown the same image filtered with a high pass kernel. This test tries to test the sensibility of the fractal techniques to this kind of noise.</p>																																				
2		<p>Multifractal spectrum is the most affected, while figure 2 spectrum shows a peak of FD which corresponds to the density value belonging to the tree threshold, the figure 1 looses the peak, since the shape is distributed along multiple density values.</p>																																				
<div><p>Efecte d'una baixa frequencia de fons</p><table border="1"><caption>Approximate data points from the graph 'Efecte d'una baixa frequencia de fons'</caption><thead><tr><th>x</th><th>1.- sense filtre (green)</th><th>2.- sumant el pas-alt (red)</th></tr></thead><tbody><tr><td>0</td><td>0.0</td><td>0.0</td></tr><tr><td>10</td><td>1.1</td><td>1.2</td></tr><tr><td>20</td><td>1.1</td><td>1.25</td></tr><tr><td>30</td><td>1.1</td><td>1.35</td></tr><tr><td>40</td><td>1.1</td><td>1.2</td></tr><tr><td>50</td><td>1.0</td><td>1.2</td></tr><tr><td>60</td><td>0.9</td><td>1.25</td></tr><tr><td>70</td><td>1.1</td><td>1.25</td></tr><tr><td>80</td><td>1.1</td><td>1.4</td></tr><tr><td>90</td><td>1.1</td><td>1.5</td></tr><tr><td>100</td><td>1.0</td><td>0.7</td></tr></tbody></table></div>			x	1.- sense filtre (green)	2.- sumant el pas-alt (red)	0	0.0	0.0	10	1.1	1.2	20	1.1	1.25	30	1.1	1.35	40	1.1	1.2	50	1.0	1.2	60	0.9	1.25	70	1.1	1.25	80	1.1	1.4	90	1.1	1.5	100	1.0	0.7
x	1.- sense filtre (green)	2.- sumant el pas-alt (red)																																				
0	0.0	0.0																																				
10	1.1	1.2																																				
20	1.1	1.25																																				
30	1.1	1.35																																				
40	1.1	1.2																																				
50	1.0	1.2																																				
60	0.9	1.25																																				
70	1.1	1.25																																				
80	1.1	1.4																																				
90	1.1	1.5																																				
100	1.0	0.7																																				
<p>Multifractal spectrum is not sensible to the grey level gradient, the</p>																																						



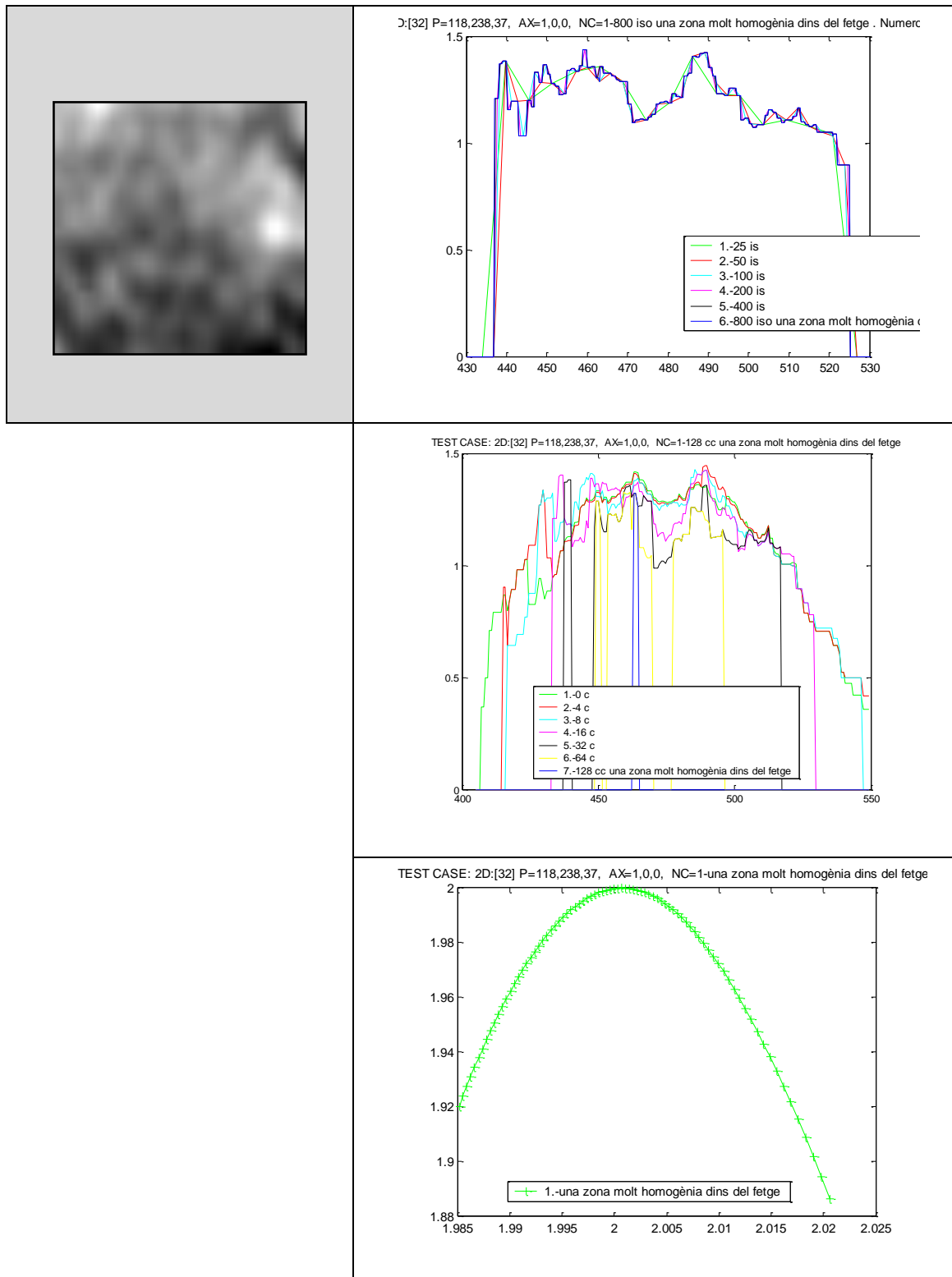
spectrum suffers non-critical variations on  $D_0$  and  $D_1$ , and a more significant decrement of the right slope. Hence it could be deduced that this approach is less sensible to low frequency noise.



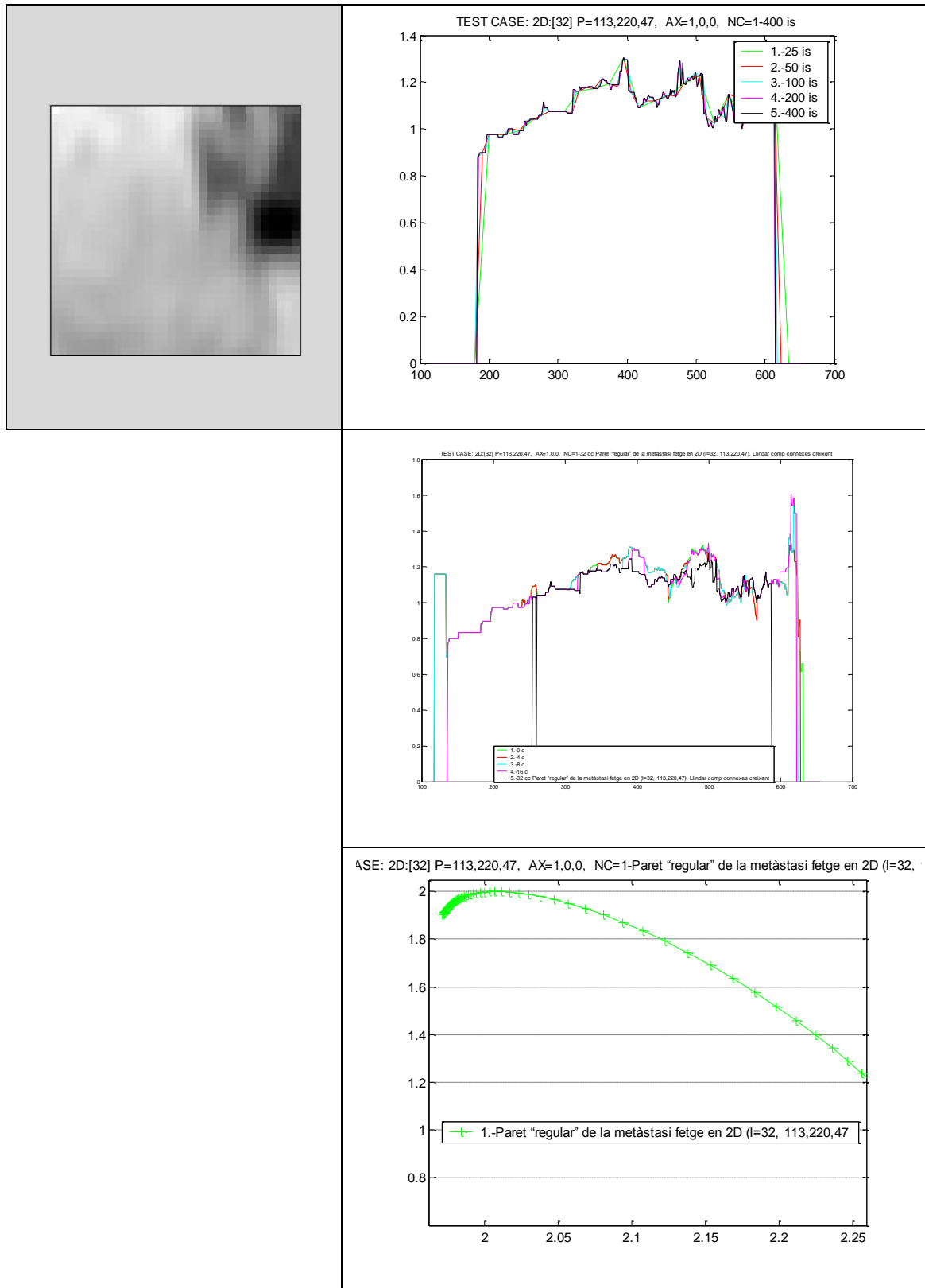
## 13.2 Medical image sample cases for multifractal and iso fractal spectrum.

Following figures show the iso fractal and multifractal spectrum of several tissues. For each sample two iso fractal figures are displayed, the first shows the evolution of spectrum when the number of sampled iso fractal levels is increased while the second studies the evolution of the spectrum when increasing the connected threshold size value. Finally, the third figure shows the multifractal spectrum.

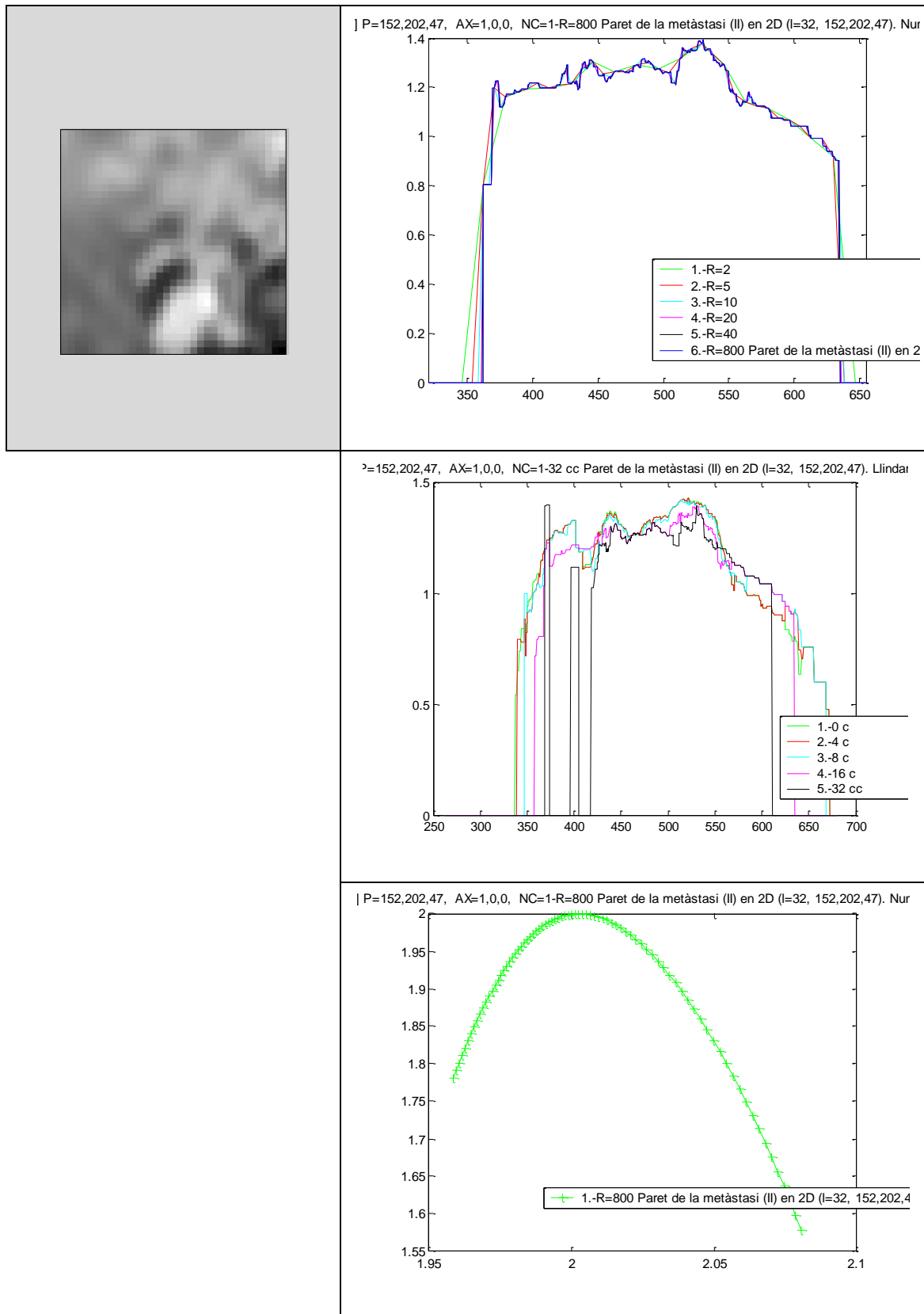
### 13.2.1 Homogeneous liver area ( $I=32$ , $\{118,238,37\}$ ).



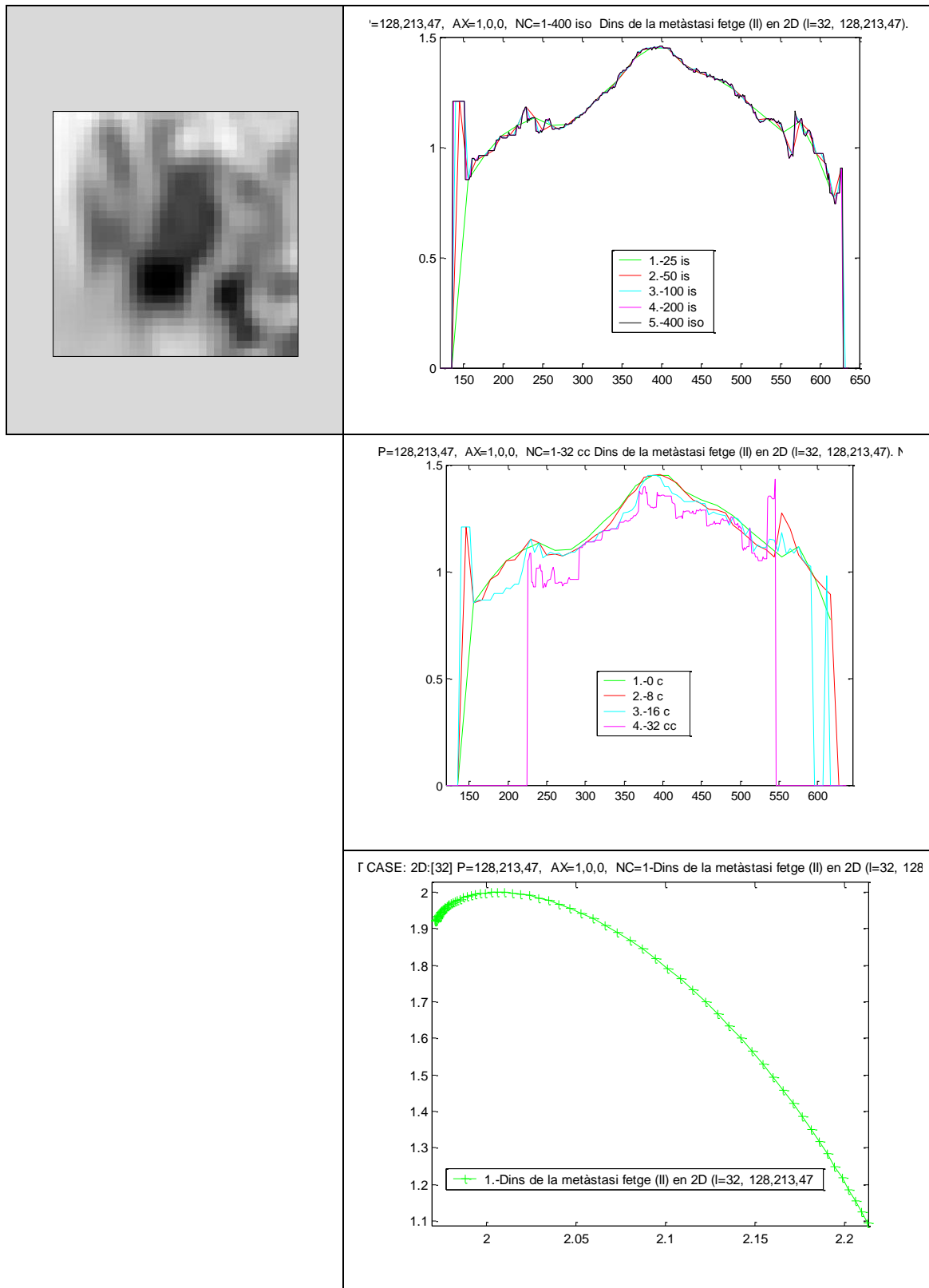
### 13.2.2 Homogeneous Tumoral border area ( $l=32$ , $\{113,220,47\}$ )



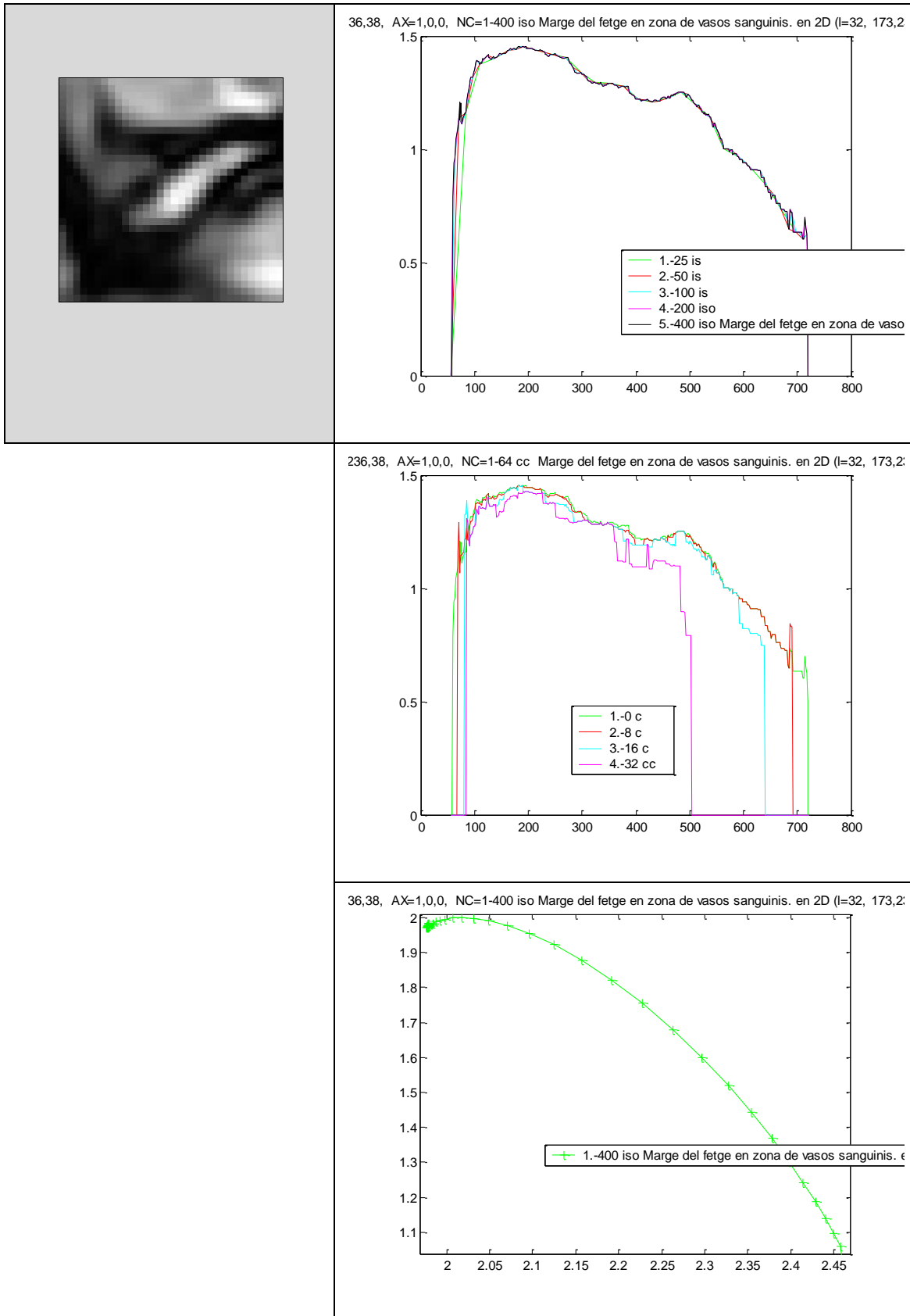
### 13.2.3 Non-Homogeneous Tumoral border area ( $l=32, \{152,202,47\}$ )



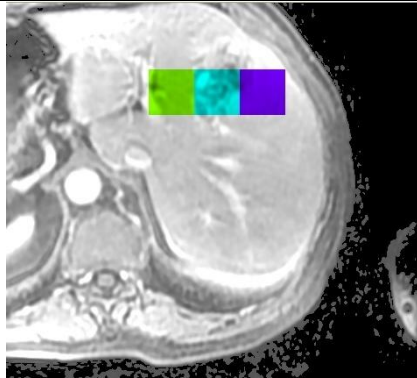
### 13.2.4 Inside the Tumoral area ( $l=32, \{128,213,47\}$ )



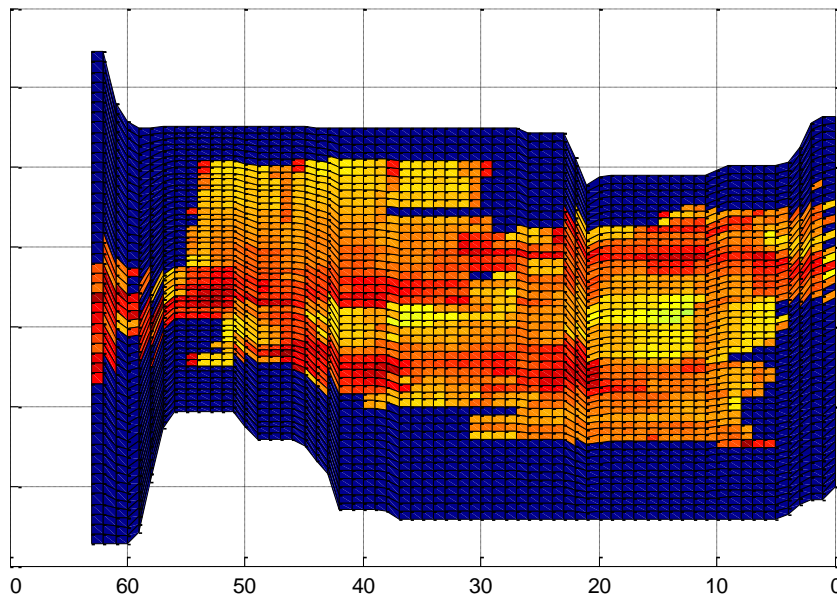
### 13.2.5 Liver border and vessels ( $l=32, \{173,236,38\}$ )



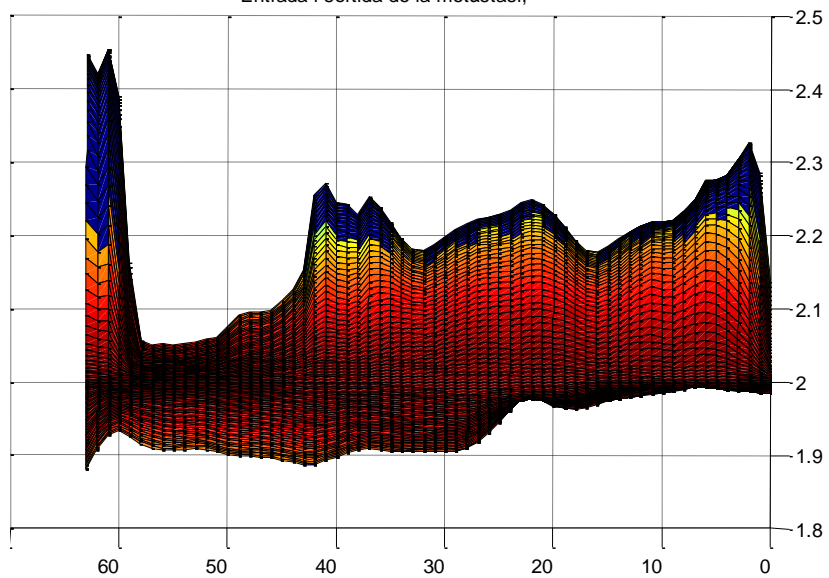
### 13.2.6 Scan-line in and out the metastasis. 64 samples path.



Entrada i sortida de la metastasi, L32)



Entrada i sortida de la metastasi,



## 13.3 User Interface

The user interface of the DimFract module is organized in four sections, each one enclosed within a group box. The module widget is shown in the left area on the figure below:

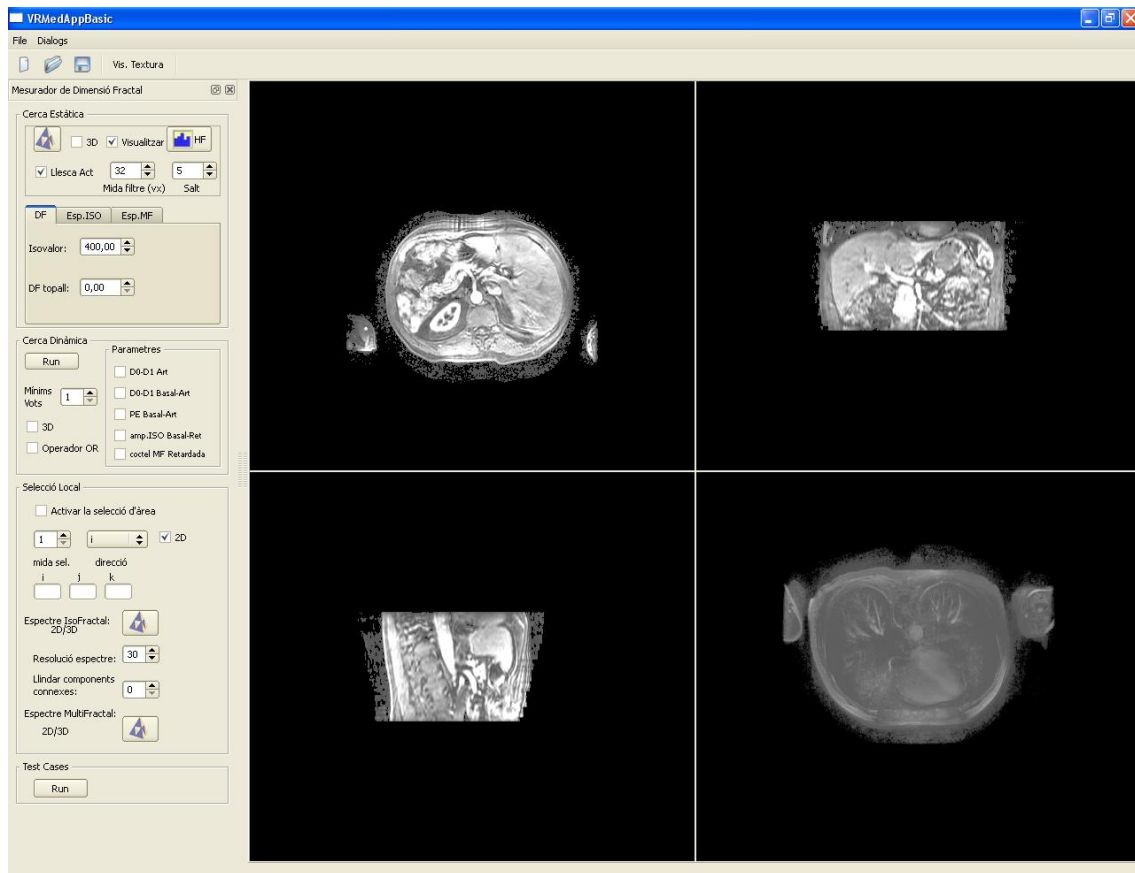


Figure 13.1: DimFract user interface widget

### 13.3.1 Static selector

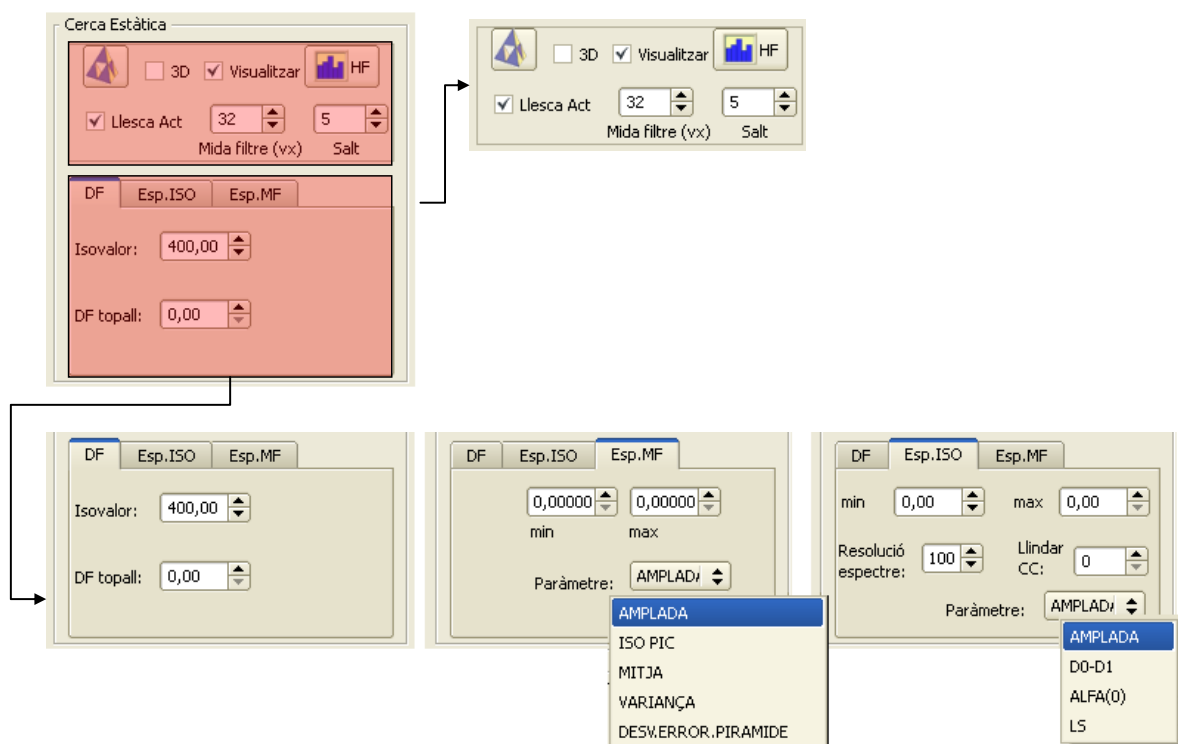
The first section is utilized for static global selection. Here the idea is visualizing the fractal dimension or any of the descriptors of the iso fractal or the multifractal spectrum. The top area is devoted to the common variables, such as window size, gap size and 3D/2D box counting. It also provides a checkbox to limit the process to the current slide.

The bottom area is a tabbed control where the user could select the algorithm to be used and adjust its specific working parameters set. Let's describe the options with more detail:

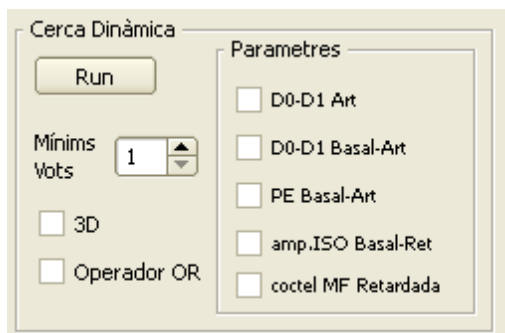
- **Fractal dimension:** A density value is required in order to select the isocurve whose fractal dimension is calculated. Additionally, a minimum threshold could be supplied in order to select only those areas above the supplied value.
- **Iso fractal spectrum:** The user must choose a descriptor that will represent the spectrum, and a minimum and a maximum value for the given descriptor in order to range the visualization. If the values are zero, the range boundaries are ignored. The following descriptors are available (their description may be found on the section 8.3):
  - Spectrum width (ISW)
  - density peak position (DP)



- spectrum mean
- spectrum variance
- Pyramid Shape Factor (PSF)
- **Multifractal spectrum:** It utilizes the same mechanism than the former, but in this case applied to multifractal spectrum. In this case the parameters are:
  - Spectrum Width (MSW)
  - $f(0)-f(1) = (D0D1)$
  - $\alpha(0)$  (A).
  - Left Slope (LS)



### 13.4 Dynamic selector



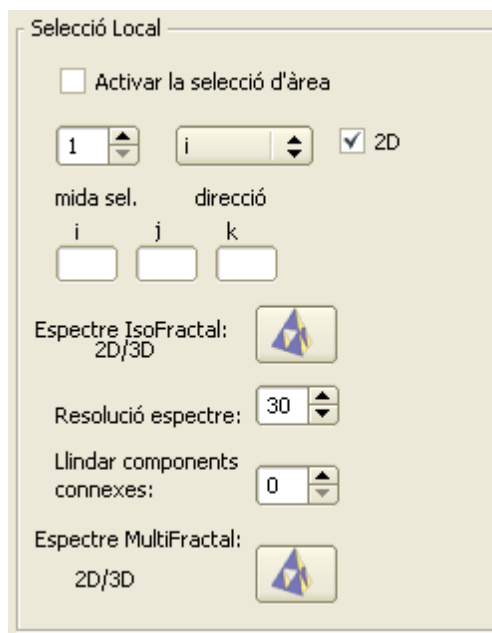
The second section covers the global search, but this time taking into account the evolution of certain parameters in time, by using the contrast sequence formed by four images (basal-arterial-venous-retarded). The values of the main iso and multifractal parameters have been precalculated and stored in files in order to accelerate the visualization process, which would be unfeasible otherwise. The user could select the parameter constraints used to select the affected areas. Each

parameter consist of a condition which is true when a given spectrum descriptor is inside a predefined range. Each windowed region that accomplish the parameter condition, receives a vote.

Parameters could be combined in order to obtain more restrictive selectors. Default combination is performed using and AND operator, but this could be changed to OR by enabling the corresponding checkbox. In addition, since the parameters are calculated in a 2D fashion, it is possible to explode the 3D spatial coherence by activating the 3D checkbox. If enabled, the selected regions extend the vote to neighbouring areas on adjacent slices.

Finally, the user can adjust the minimum number of votes required for an area to appear as active. The default value is 1.

### 13.5 Local selection



Local selection utilities group is used mainly to locate, select and study certain areas of the volume. It permits window placing by interactively moving the selected area with the mouse, or alternatively the user may specify directly the *ijk* window coordinates. The shape and dimensions of the window are determined by the *size* and *direction* variables. If *size* is one, the analysis window is a square of size *W*. If *size* is greater than 1, the analysis windows is a parallelepiped of  $W \times W \cdot size$ . The parallelepiped longer side direction is determined by the axis selected, which could be switch to *i*, *j* or *k*. If the user deselects the 2D checkbox, the program will select cubes instead of boxes.

Once the selection cursor highlights the desired area, the user could launch the calculation of the iso fractal or multifractal spectrum using the corresponding buttons. This action will generate a

matlab file called *graph.m*, which contains the spectrum data and plots a graph showing up the spectra.

### 13.6 Test cases



Finally, the last option permits launching the test cases file. The program expects the existence of the file *./testCases/index.test* in the execution path. If it is found, it will execute the instructions contained

in the file. Finally, the application will store the results in the *graph.m* file.

*index.test* is a plain text ASCII file. Each line contains the description of a test case, and all the parameter set required for its calculation. The fields inside the test-case line are tab-separated values following a closed specification. Each line starts with a test-case code type.

There are two main categories of test cases, those applied to a specific synthetic image, which must supplied in the shape of a raw format file, and those applied to the volumetric image which has been already open in the application. The next table shows the test-case types available, as well as their codes.

Type	Code	Test-case description
Synthetic image ( images must be squares with side $n^2+1$ )	0	Isofractal analysis applied to the whole image
	1	Multifractal analysis applied to the whole image
Volumetric image ( applied to the current image )	2	Isofractal analysis. Defines an iso fractal analysis on a user-defined window.
	3	Multifractal analysis. Defines a multifractal analysis on a user-defined window.
	4	Crossed-Parameter analysis. Given a pair of specific descriptor codes and an area, the application generates a plot, using each descriptor as x and y coordinates respectively, and placing a marker in the coordinates corresponding to the specified area.

The test cases could be easily grouped in a same plot by indicating an adequate total and serial number in the line (fields [TTC] and [#TC]). The specification of the fields' descriptions for each test-case type is detailed below:

0	[TTC]: Total of test cases in the serie. (1 if it is a single test case)	[#TC]: Number of test case in the serie (1 if it is a single test case)	Image path	Side pixel size ( $n^2+1$ )	[NS]:Isofractal spectrum number of sample levels	[MCCT]: Minimum connected components threshold.
---	--	---	------------	-----------------------------	--	---

1	[TTC]	[#TC]	Image path	Side pixel size ( $n^2+1$ )		
---	-------	-------	------------	-----------------------------	--	--

Window Coordinates

2	[TTC]	[#TC]	[D]:Dimension [ 2   3 ]	[WS]: Window Size ( $n^2$ )	[NS]	I	J	k	[AD]:Axis direction [ i   j   k ]	[NBAD]:Number of boxes in the axis direction ( 1 if square )	[MCCT]	[CM]: Comments
---	-------	-------	----------------------------	-----------------------------	------	---	---	---	--------------------------------------	--	--------	----------------

3	[TTC]	[#TC]	[D]	[WS]		I	J	k	[AD]	[NBAD]		[CM]
---	-------	-------	-----	------	--	---	---	---	------	--------	--	------

4	[TTC]	[#TC]	[D]	[WS]	[NS]	I	J	k	[AD]	[NBAD]	[MCCT]	Parameter x code	Parameter y code	DF minimum threshold for isospectrum width calculation.	[CM]
---	-------	-------	-----	------	------	---	---	---	------	--------	--------	------------------	------------------	---	------

The descriptor codes for type 4 are selected from the following table:

0	Density peak position (DP)
1	General IsoSpectrum width (ISW)
2	<i>Not in use</i>
3	alfa(0) (A)
4	f(0)-f(1) (D0D1)
5	Multi spectrum width (MSW)
6	IsoSpectrum width above a certain FD threshold (ISW)
7	Multi spectrum Horizontal balance (HB).
8	Multi spectrum vertical balance (VB).
9	Left Slope (LS)
10	Pyramid Shape Factor (PSF)

## 13.7 DimFract module Class Diagram

