# Nuclear imaging analysis in a multimodal and multitemporal context

Francisco Javier Alvarez Padilla

UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

ÉCOLE DOCTORALE SCIENCES DU NUMÉRIQUE ET DE L'INGÉNIEUR (SNI) N°620

# THÈSE

Pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE**

***Discipline*** : ***INFORMATIQUE***

Présentée et soutenue publiquement par


## FRANCISCO JAVIER ALVAREZ PADILLA


le 13 septembre 2019

---

## Analyse d'images nucléaires dans un contexte multimodal et multitemporel

---

Thèse dirigée par **DIMITRI PAPATHANASSIOU**,

co-dirigée par **NICOLAS PASSAT**,

et encadrée par **BARBARA ROMANIUK**

**JURY**

| | | | |
|---|---|---|---|
| M. Bertrand KERAUTRET | Professeur des Universités | Université Louis Lumière Lyon 2 | **Rapporteur** |
| M. Antoine VACAVANT | Maître de Conférences | Université Clermont Auvergne | **Rapporteur** |
| M. Hugues TALBOT | Professeur | École Centrale Paris | **Examinateur** |
| Mme Barbara ROMANIUK | Maître de Conférences | Université de Reims Champagne-Ardenne | **Examinateur** |
| Mme Stéphanie SERVAGI-VERNAT | Maître de Conférences - Praticien Hospitalier | Institut Jean Godinot | **Examinateur** |
| M. Dimitri PAPATHANASSIOU | Professeur des Universités - Praticien Hospitalier | Université de Reims Champagne-Ardenne | **Directeur** |
| M. Nicolas PASSAT | Professeur des Universités | Université de Reims Champagne-Ardenne | **Co-directeur** |


UNIVERSITÉ
DE REIMS
CHAMPAGNE-ARDENNE


UNIVERSIDAD DE GUADALAJARA
PIENSA Y
TRABAJA


CReSTIC


Institut Godinot
unicancer Reims en Champagne

# Dedication

This work has been a big challenge to my body and my mind. Today I am happy because of finishing what I started as better as I could do.

I dedicate my dissertation to my family,

> my parents Francisco Javier ALVAREZ CHAVEZ and Sandra Cristina PADILLA ZAMORA, who motivate me all the time to follow this academic road, since they "force me" to learn French language until they never cut down money in health and education if it was necessary;
>
> my twin brother Juan Rodolfo ALVAREZ PADILLA, with whom I grew up elbow to elbow encouraging each other, now by discussing of his new professional projects on control theory in the academia and in enterprise;
>
> the big baby Cristian ALVAREZ PADILLA, the filmmaker of the family, who cheers me up with his new interesting film projects;
>
> my aunt Maria del Carmen PADILLA ZAMORA as my second mother, whose cooking classes helped me survive far from home;
>
> and my grand mothers Maria Luisa ZAMORA ESPINOSA and Maria Angelina CHAVEZ DEL TORO, who start this goal at my side but life take them away in physic but staying in my memory;

to them, by all their words and actions of encouragement that support me throughout the entire doctorate program.

# Acknowledgments

I firstly thank my body, that in spite of several "medical unforeseen", has endured until this day.

I thank Antoine VACAVANT and Bertrand KERAUTRET as thesis reporters; Hugues TALBOT and Stéphanie SERVAGI-VERNAT as examiners; Benoît NAEGEL as my guest; and Nicolas PASSAT, Dimitri PAPATHANASSIOU and Barbara ROMANIUK as my research group; who together are part of the referees for assessing this work.

I wish to thank all my thesis group, leaded by my directors Nicolas PASSAT on exact sciences support and Dimitri PAPATHANASSIOU on medical application field, particularly Nicolas for his several strongly argued corrections showing his interest for my formation as formal researcher; my local supervisor Barbara ROMANIUK for her advices in the following of my thesis, in terms of organization and coherence in writing expression; Benoît NAEGEL, as external supervisor, for his guidance and correction modeling the methodological propositions that were developed in my work; and Stéphanie SERVAGI-VERNAT for her contribution providing and guiding a medical case (problem/data).

I thank my PhD colleagues, David MORLAND for his medical context assistance and provision of majority of my medical image sets; Joel RANDRIANANDRASANA, Nicolas COURILLEAU, Jonathan SARTON and Ludovic BLACHE for their programing advices; and Ulysse LARVY DELARIVIERE, Joris FALIP, Pierre CETTOUR-JANET, Jimmy Francky RANDRIANASOA for their time at feedback discussions.

I thank Edwin CARLINET for providing to this work the pylene library and his assistance on it, for building and processing of a component-tree and a tree of shapes.

I thank Sofiane GUENDOUZEN for his guidance, as medical physicist, to produce exploitable files for clinical application.

I thank Barbara ROMANIUK and Nicolas PASSAT for being my guarantors which is an obligatory requirement, under my foreign incomes, to apply for a location in France.

I thank Joel RANDRIANANDRASANA and Pierre CETTOUR-JANET to host me at their home when I lived some homeless (*SDF*) lapses.

I thank my twin brother Juan Rodolfo ALVAREZ PADILLA, for the time he spent on me, discussing about mathematics and some ideas of my thesis methods I wanted to clarify by talking with someone. He was often there to listen to me.

Finally, I thank my Colombian friends, Erick ARIZA and Juan HERRERA, for their moral support at Reims, making me feel, by latino culture, that I was in my Mexico, with my people, for some precious instances.

# Introduction

## Contexte de la thèse

Le cancer est un problème d'actualité de nos jours en raison de 17 millions de cas dans le monde avec un taux de mortalité de 56,4%, selon *Globalcancer statistics 2018*. La chimie, la pharmacobiologie et l'imagerie médicale sont quelques-uns des domaines de recherche qui ont porté leur attention sur ce sujet afin d'examiner les solutions respectives et les développements permettant de pallier conjointement les problématiques oncologiques. Dans le domaine de l'imagerie médicale, la Tomographie par Émission de Positons (TEP) / Tomodensitométrie (TDM) est largement utilisée pour les applications en oncologie afin de fournir des données hautement pertinentes *in vivo* sur les tissus du corps humain : le niveau métabolique et la densité tissulaire. Ces informations peuvent fournir de riches pistes pour détecter la présence et la position de tumeurs dans le corps. Les experts médicaux exploitent les données de TEP / TDM pour segmenter manuellement les tumeurs à l'aide d'outils interactifs basés sur une visualisation fusionnée. Cependant, ce mode de fonctionnement est chronophage et peut entraîner une variabilité intra et inter-opérateur.

## Motivation de la thèse

Le besoin d'outils efficaces pour traiter automatiquement les images médicales dans la pratique clinique, en gagnant du temps et de la qualité pour favoriser le traitement du patient, a conduit à proposer plusieurs méthodologies. Ces méthodologies s'inspirent de différents domaines et essayent de surmonter les inconvénients intrinsèques de l'image et d'améliorer la précision de la segmentation.

Récemment, des approches multimodales TEP / TDM ont été proposées, en imitant les experts médicaux en application clinique pour tirer parti des spécificités anatomiques / fonctionnelles pour la segmentation de la tumeur.

Suivant cette tendance, nous cherchons à développer des stratégies méthodologiques en nous focalisant sur la segmentation tumorale.

## Contributions

Ce travail de thèse est méthodologiquement situé à l'interface entre l'analyse d'image et le traitement d'image bas niveau et ayant des images médicales 3D

TEP / TDM comme source principale d'information ainsi que les tumeurs comme étant nos objets d'intérêt à des fins de segmentation.

Trois méthodologies sont développées dans ce travail, partageant l'utilisation d'une structure de données arborescente, mais différant dans la manière de traiter les informations provenant de telles structures hiérarchiques.

Dans un premier travail, un arbre est utilisé comme extracteur de données pour être injecté dans une méthode de marcheur aléatoire pour la segmentation.

Dans un deuxième travail, notre approche propose une solution de gestion de la dépendance spatiale par comparaison nœud à nœud sur des arbres pour une procédure de mise en correspondance sur le critère de similarité des nœuds suivie d'une phase de segmentation.

Dans un troisième travail, une arborescence est utilisée pour traiter et segmenter les nœuds adaptant le randonneur aléatoire à leur structure hiérarchique et ceci afin d'exploiter pleinement les avantages de cette structure arborescente.

## Contexte administratif et contexte collaboratif

Ce travail a été financé par une bourse doctorale, provenant de l'*Universidad de Guadalajara* au Mexique et partiellement financé par "le Programme Investissement d'Avenir" de l'Agence Nationale de la Recherche ("Infrastructures d'Avenir en Biologie Santé", ANR-11-INBS-0006).

La direction a été confiée au laboratoire CReSTIC, unité de recherche, rattachée à l'Université de Reims Champagne-Ardenne en tant qu'institution académique; et co-dirigé par l'Institut Jean Godinot comme établissement clinique.

Deux collaborations sont ici à mentionner. Un premier partenariat avec le laboratoire ICube a consisté en une supervision externe assurée par Benoit Naegel en raison de ses compétences dans le domaine du traitement d'images et tout particulièrement sur les arbres de coupes. Un second partenariat avec le Dr. Stéphanie Servagi-Vernat au centre Jean Godinot a consisté à segmenter les tumeurs de la tête et du cou pour une sous-segmentation en fonction de l'activité hétérogène de la tumeur et ainsi planifier la radiothérapie à intensité modulée.

## Organisation du manuscrit

Ce manuscrit est composé de 3 chapitres organisés comme suit :

**Chapitre 1 :** Nous présentons ici une introduction au problème sur lequel porte cette thèse, en ce qui concerne le contexte clinique et les limitations techniques. Trois sections sont utilisées à cette fin : d'abord, le **contexte médical** pour présenter des statistiques sensibles sur le cancer dans le monde entier, son incidence et la mortalité, en soulignant la pertinence de la recherche sur ce sujet. Les images médicales, notamment TEP / TDM, sont présentées comme un outil largement utilisé pour l'extraction de données à des fins de diagnostic et de planification du traitement de plusieurs types de cancer. Une introduction à l'**imagerie TEP / TDM** est proposée

pour expliquer leurs spécificités tant pour la détection de l'activité tumorale que l'identification de son contexte anatomique, en se basant sur leurs caractéristiques monomodales, leur acquisition et leur reconstruction. Méthodologiquement partant, un **état de l'art pour la segmentation TEP** est présenté, en exposant les travaux précédemment proposés par les auteurs sur ce domaine et en identifiant les contraintes auxquelles nous essayons de faire face.

**Chapitre 2 :** La **proposition méthodologique pour la segmentation** repose sur des aspects communs, principalement axés sur le traitement des structures hiérarchiques, qui seront introduits dans la première section **fondements théoriques**, suivie de la section **méthodes développées**, présentant les particularités des modèles algorithmiques que nous avons proposés.

**Chapitre 3 :** Une **instanciation méthodologique en TEP / TDM** est ici présentée. Une section sur les **paramètres généraux** sur les méthodes développées, est introduite pour la segmentation tumorale par TEP, TDM et TDM avec injection de contraste. Les expérimentations sur les **méthode 1**, **méthode 2** et **méthode 3**, les résultats et les conclusions y sont finalement présentés.

**Conclusion :** Un dernier chapitre est consacré à la présentation des **conclusions et perspectives** afin d'évaluer nos résultats inter-méthodes et aboutir à des conclusions globales. Nous soulevons ici l'impact médical que peuvent avoir nos résultats de segmentation sur la planification en radiothérapie. Finalement nous présentons les perspectives pour nos travaux futurs.

# Introduction

## Thesis context

Cancer disease is a highly sensitive problem nowadays due to 17 million cases worldwide with a 56.4% mortality rate, according to global cancer statistics 2018. Chemistry, pharmacobiology and medical imaging are some of research fields that have paid their attention on this topic to investigate respective solutions an developments to jointly palliate these oncological issues. Within the field of medical imaging, Positron Emission Tomography coupled with Computed Tomography is widely used for oncology applications to provide highly relevant data in vivo of human body tissues: metabolic level and composition density. This information can provide rich tracks to detect the presence and position of tumors in the body. Medical experts exploit PET/CT data to manually delineate tumors with interactive tools based on a fused visualization. However, this handmade modus operandi is time-consuming and can lead to intra- and inter-operator variability.

## Thesis motivation

The need for efficient tools to automatically process medical images in clinical practice, gaining time and quality to favor patient treatment, has led to the proposal of several methodologies, from different domains, trying to overcome intrinsic image drawbacks and improve segmentation accuracy.

Recently, PET/CT multimodal approaches have been proposed, emulating medical experts in clinical application to take advantage of anatomical/functional specificities for tumor segmentation.

Following this tendency, we developed methodological strategies that are focused on this tumor segmentation framework.

## Contributions

This thesis work is methodologically located at the interface between image analysis and low-level image processing; having 3D PET/CT medical images as main information source and tumors as our objects of interest for segmentation purposes.

Three methodologies are developed in this work, sharing the use of a tree data-structure, but differing in the way of information processing from such hierarchical objects.

In a first work, a tree is used as data extractor to be injected into a random walker method for segmentation.

In a second work, our approach proposes a solution to handle spatial dependency by node-to-node comparison on trees for a matching procedure on their similarity, followed by a segmentation step.

In a third work, a tree is used to process and segment nodes adapting the random walker on their hierarchical structure, to fully exploit the tree advantages.

## Thesis administrative and collaboration context

This work was funded by a PhD grant from *Universidad de Guadalajara*, México; and partly funded by the *Programme Investissement d'Avenir* of the French *Agence Nationale de la Recherche* ("*Infrastructures d'Avenir en Biologie Santé*", ANR-11-INBS-0006).

Direction is assumed by the CReSTIC laboratory as scientific organism, attached to *Université de Reims Champagne-Ardenne* as academic institution; and co-directed by *Institute Jean Godinot* as clinical institute.

Two collaborations are here added, a partnership with ICube laboratory, consisting of an external supervision provided by Benoit Naegel due to his skills in the field of image processing and component-tree specialization; and a second collaboration with Dr. Stephanie Servagi-Vernat, consisting of segmentation of head and neck tumors for a sub-segmentation according to heterogeneous activity and planning of intensity-modulated radiation therapy.

## Manuscript organization

This manuscript is composed of 3 chapters that are organized as follows:

**Chapter 1:** An introduction to the problem that this thesis is focused on, with regards to clinical background and technical limitations, is made in this chapter. Three sections are used to this end. **Medical context** to present sensitive cancer statistics around the world about incidence and mortality, underlining the research relevance on this topic. Medical images, particularly PET/CT, are presented as a widely used tool for data extraction for purposes of diagnosis and treatment planning for several cancer types. A PET/CT overview, from **PET/CT origins** and **PET/CT fundamentals**, is made to explain their specificities for tumor activity and anatomy evaluation, their monomodal characteristics, acquisition and reconstruction; and the interest obtained with their multimodal sequential acquisition. Finally, a **state of the art for PET segmentation** is presented, exposing the previous works that authors have proposed on this domain and defining the drawbacks we attempt to face.

**Chapter 2:** Our **methodological proposals for segmentation** share some common aspects, mainly focused on hierarchical structure processing, that are going to be introduced in **theoretical basis** first section, followed of **developed segmentation methodologies** section, presenting algorithmic model particularities.

**Chapter 3:** The **methodological instantiation in PET/CT** is here presented. A section on **general parameters** of the developed methods is introduced for PET, CT and contrast-enhanced CT tumor segmentation. **Method 1, method 2** and **method 3** experiments, results and conclusions are shown in followed sections.

**Conclusion:** A final **conclusions and perspectives** chapter is dedicated to make a research summary to evaluate our inter method results and global conclusions; share medical application of our segmentation results for radiotherapy planning; and finally talk about the perspectives in future works.

# Contents

# List of Tables

# List of Figures

# Glossary

| Notation | Description | Page List |
|---|---|---|
| $A$ | Vector of region-based attributes | 64 |
| $I$ | Matrix image | 54 |
| $I_A$ | Vectorial image from $\Psi$ | 64 |
| $N$ | Node of a tree $\Psi$ | 54 |
| $N_c$ | Operator to map to all node children $N_c$ of $N_i$ | 55 |
| $N_p$ | Operator to map to node parent $N_p$ of $N_i$ | 55 |
| $Q$ | Support of intensity levels | 54 |
| $\Delta$ | $\pm$ node interval | 65 |
| $\Omega$ | Support of pixels/voxels | 54 |
| $\Theta$ | Support of nodes | 54 |
| $\beta$ | Scalar or vectorial coefficients to weight respective metric influence at random walker method | 62, 69 |
| $\beta^+$ | Positron particle disintegrated from atom nucleus | 34 |
| $\beta^-$ | Electron particle disintegrated from atom nucleus | 34 |
| $\hat{N}$ | Pruned node of $\Psi$ | 65 |
| $\hat{\Psi}$ | Reconfigured pruned $\Psi$ | 65 |
| $\phi$ | Tree of shapes | 56 |
| $\tau$ | Component-tree | 55 |
| $\varphi$ | Nested node set at $\Psi$ from a node $N_i$ to a node $N_j \in N_{anc}(N_i)$ | 55 |
| $\vartheta$ | Intensity $q$ related to the partitioning condition for node $N$ conception | 55 |
| $n_c$ | Operator that gives the number of $N_c$ of $N_i$ | 55 |
| $n_p$ | Operator that gives the number of $N_p$ of $N_i$ | 55 |
| $o$ | Connected-component/object | 95 |
| $p$ | Probability map as random walker method result | 63 |
| $q$ | Intensity level | 54 |
| $v$ | Node of a graph $G$ | 62 |
| $w$ | Weight of edge $e_{ij}$ | 62 |
| $x$ | Pixel/Voxel | 54 |

# Acronyms

| Notation | Description | Page List |
|---|---|---|
| $SE$ | Structuring element | 95 |
| $bg$ | Background | 89 |
| $fg$ | Foreground | 89 |
| CE-CT | Contrast-Enhanced Computed Tomography | 86 |
| CT | Computed Tomography | 24 |
| DICOM | Digital Imaging and COmmunications in Medicine | 33 |
| FDG | Fluorodeoxyglucose | 26 |
| FFT | Fast Fourier Transform | 30 |
| fMRI | fonctional Magnetic Resonance Imaging | 24 |
| HU | Hounsfield Units | 88 |
| kBq | KiloBecquerel; radionuclide activity measure unit | 33 |
| keV | KiloElectronVolt | 28 |
| LOR | Line Of Response represents two detected incident photons from same anhilation event | 28 |
| MIP | Maximum Intensity Projection | 41 |
| MRI | Magnetic Resonance Imaging | 24 |
| MSER | Maximally Stable Extremal Region | 61 |
| PET | Positron Emission Tomography | 23 |
| PVE | Partial Volume Effect | 27 |
| RW | Random Walker method | 62 |
| SCont | Spatial Contrast | 89 |
| SPECT | Single-Photon Emission Computed Tomography | 24 |

| Notation | Description | Page List |
|---|---|---|
| SUV | Standard Uptake Value | 33 |
| SVM | Support Vector Machine | 40 |
| TOF | Time Of Flight | 29 |

# Chapitre 1 : Contexte de thèse

## Contexte médical

Malgré des avancées cliniques significatives dans la recherche et le développement de dispositifs technologiques, le cancer est toujours un problème de santé majeur dans le monde, dont le nombre de décès est estimé à 9,5 millions en 2018 selon *Global cancer statistics* [1]. Ces statistiques ont accru l'intérêt de nombreux centres de recherche, dans les domaines dédiés à l'oncologie, comme la biologie, la chimie, et l'informatique par exemple, pour trouver des améliorations et des solutions en faveur de la santé de l'être humain. En informatique, les images médicales, depuis leur reconstruction jusqu'à leur traitement, ont été et sont toujours d'un grand intérêt. Ceci a permis de développer des outils performants au service des personnels cliniques, dont la segmentation automatique des tumeurs est un exemple emblématique.

Intrinsèquement, les images médicales permettent d'extraire des informations *in vivo* du patient, de manière non-invasive, pour le diagnostic, la planification des traitements et le suivi du patient. La Tomographie par Émission de Positons (TEP) est largement utilisée en oncologie pour l'analyse du métabolisme cellulaire où les tumeurs malignes sont caractérisées pour leur croissance agressive et donc par un fort métabolisme. La tomodensitométrie (TDM), ou carte de l'atténuation des rayons X, est régulièrement acquise séquentiellement avec la TEP à des fins de correction de l'atténuation anatomique (nature physique de la TDM) et de co-visualisation anatomique et fonctionnelle du patient.

Le couplage TEP / TDM, les caractéristiques de chacune des deux modalités ainsi que l'intérêt d'une telle acquisition bimodale permettent de définir les enjeux qui peuvent être abordés à travers ce type d'examens.

## Imagerie TEP / TDM

La TEP propose une représentation 3D d'un métabolisme, produite par la détection des photons gamma coïncidents, provenant de l'annihilation électron-positon. En fonction de l'évaluation du métabolisme, un biomarqueur lié à un radionucléide à décroissance $\beta^+$, dont le glucose radioactif FDG est le plus utilisé, est sélectionné pour être injecté au patient. Ceci provoque un processus physico-chimique au sein du corps de celui-ci. Le biomarqueur est guidé à travers le réseau de vaisseaux par le flux sanguin, se fixant dans les cellules selon leur métabolisme (ou étant évacué du corps).

Les photons peuvent être sensibles aux phénomènes ionisants de la matière lorsqu'ils traversent la masse corporelle en direction des capteurs. En particulier,

ils sont sensibles aux phénomènes de diffusion (effet Compton) et d'atténuation. La carte d'atténuation obtenue depuis la TDM, principale raison de cette acquisition bimodale, est donc utilisée pour pallier l'atténuation dans la TEP. Cette acquisition séquentielle permet d'obtenir une correspondance spatiale approximative qui peut être exploitée pour une co-visualisation fonctionnelle avec des références anatomiques.

En tant que modalités tomographiques, la TEP et la TDM sont soumises à une reconstruction basée sur la transformée de Radon. Pour la TEP spécifiquement, cette reconstruction tomographique est fréquemment réalisée à partir de l'algorithme de calcul direct par rétroprojection filtrée ou l'algorithme itératif OSEM.

La TEP permet d'analyser les informations sur la fixation locale du radiotraceur, dont la signification des valeurs est directement dépendante de la dose totale injectée. La valeur de fixation normalisée SUV (*Standard Uptake Value*) a été proposée comme mesure de référence des intensités de la TEP par rapport à la distribution de la radioactivité du radiotraceur injecté à travers le corps. Elle permet notamment de déterminer la différence entre les niveaux de fixation normaux et anormaux.

Un radiotraceur ou un biomarqueur est une substance radioactive utilisée pour marquer une activité physiologique. Elle est obtenue en associant un isotope radioactif à une molécule d'une substance telle que le glucose dans le cas du FDG. En oncologie, les radiotraceurs permettent de mettre en évidence des activités métaboliques anormales, caractéristiques des cellules cancérogènes. D'autre part, un radionucléide est le noyau radioactif d'un biomarqueur. Chimiquement, il s'agit d'un atome d'isotope dont le noyau est instable et ainsi susceptible d'émettre des particules / énergie.

Les données provenant de la TEP / TDM sont alors exploitées pour segmenter de zones tumorales chez le patient.

## État de l'art pour la segmentation TEP

L'analyse des tumeurs est essentielle pour le diagnostic, la planification du traitement et le suivi temporel des patients atteints des maladies cancéreuses. La visualisation bimodale TEP / TDM améliore la visibilité des tumeurs pour une segmentation manuelle assistée par ordinateur. Cette segmentation interactive et intuitive est communément utilisée pour définir des régions d'intérêt (zones composites, organes ou tumeurs) en vue de l'extraction de leurs caractéristiques et à posteriori la planification de la radiothérapie. Cependant, des erreurs intra- et inter-opérateur, ainsi que des problématiques intrinsèques à la TEP, comme l'effet de volume partiel ou sa faible résolution spatiale, présentent un frein à la définition précise du volume tumoral. Plusieurs méthodologies ont été développées pour prendre en compte ces difficultés. Elles visent à automatiser de manière optimale cette opération, en visant à trouver des résultats reproductibles et précis, tout en réduisant le temps de traitement.

Des nombreuses approches de segmentation monomodale TEP ont été développées, en tirant parti du contraste élevé des tumeurs. Ces méthodologies peuvent être classées en fonction du domaine de développement : approches basées sur des seuillages, basées contours, stochastiques ou encore à base des graphes.

Récemment, des approches multimodales ont été proposées avec l'objectif d'augmenter la précision de la segmentation en exploitant différentes sources de données, principalement des sources fonctionnelles et anatomiques conjointement.

L'ensemble des méthodes précédemment mentionnées sera introduit dans les paragraphes suivants afin de mettre en évidence les avantages et les limitations de méthode proposées dans l'état de l'art dans le contexte de la segmentation tumorale.

Les **approaches par seuillage** sont basées sur la sélection d'une valeur dans l'histogramme d'intensités, dont les voxels de plus haute valeur sont considérés correspondant aux zones hyperfixantes (dont la tumeur) et les autres comme le fond, convertissant ainsi une image en niveaux de gris en une image binaire. Selon la manière dont est définie la valeur du seuil optimal, ces méthodes peuvents se diviser en celles à seuillage fixe, celles à seuillage adaptatif et finalement celles à seuillage iteratif. Le seuillage fixe est défini de manière intuitive en observant l'histogramme des intensités ou par un expert. Pour tenter d'adapter la sélection des seuils en fonction des conditions d'entrée, plusieurs approches analytiques et itératives ont été présentées. Spécifiquement les approches par seuillage itératif ont été proposées pour augmenter la robustesse grace à une convergence vers un seuil optimal.

Les **approches par contours** visent à identifier les zones discontinues d'une image qui correspondent aux contours. Pour la TEP, cette tâche n'est pas évidente à cause des contours flous dus à la faible résolution spatiale. Cependant, certaines approches font appel à des techniques de prétraitement des images afin de prendre en compte directement les incertitudes. Elles peuvent être regroupées en deux catégories : les méthodes basées sur les gradients ou les modèles déformables. Un gradient est centré sur la définition de ses zones limites, généralement caractérisées par un changement d'intensité brusque. Le concept de modèles déformables consiste, généralement, à déformer un modèle initial vers une modélisation stable appropriée. Les contours actifs sont populaires dans ce domaine. Ils sont basés sur l'évolution d'un contour dynamique, qui à partir d'un état initial, va évoluer par des mouvements anisotropes (de manière analogue au déplacement d'un serpent), jusqu'à atteindre la convergence d'une fonction d'énergie. Cette énergie est composée de deux termes antagonistes : les énergies internes et externes. Ces paramètres permettent de pousser et de tirer la limite de l'objet d'intérêt à travers son environnement. L'énergie interne permet de contrôler la géométrie du modèle (force de lissage), tandis que le mouvement externe est contrôlé par les caractéristiques d'intérêt.

Les **approches par classification** visent à étiqueter chaque voxel de l'image en l'affectant à une classe. Le processus de classification est divisé en trois étapes : définition des paramètres caractéristiques de chaque voxel[1] (intensité, par exemple) ; définition des paramètres caractéristiques de chaque classe ; définition des règles de décision pour déterminer l'appartenance d'un voxel à une classe. Il existe deux familles d'approches de classification selon l'intervention éventuelle d'un opérateur : les approches supervisées et les approches non supervisées (communément appelées clustering). Dans les approches supervisées, les car-

[1] Le terme voxels est ici utilisé car les modèles sont adaptés pour travailler les images 3D. Ce formalisme reste cependant valable en 2D.

actéristiques des classes sont connues. Dans le cas d'approches non supervisées, il n'y a pas d'information *a priori* pour exécuter l'algorithme de classification. Par conséquent, une étape d'apprentissage est effectuée automatiquement par un déplacement itératif des centroïdes en fonction des membres de leur classe. Les algorithmes K-moyennes, *fuzzy-c-means* et *SVM* (séparateurs à vaste marge) sont quelques exemples d'approches pouvant être envisagées.

Les **approches par graphes** consistent généralement à projeter la matrice de voxels d'une image dans un espace de graphes structuré par des nœuds et des arêtes, en fournissant ainsi un modèle plus flexible pour gérer les intensités et les interactions entre les voxels et les régions. Le type le plus simple de graphe associé à une image $G(V,E)$ consiste à modéliser directement un voxel par un nœud V et à lier les nœuds voisins par des arêtes $E$. Dans les approches de marcheur aléatoire, ce type de graphe a été exploité en pondérant les arêtes afin de trouver le chemin le plus court entre les voxels marqués et les voxels non marqués. D'autre part, des structures hiérarchiques ont été également proposées, notamment pour modéliser une image sous forme d'un arbre enraciné, dont la racine correspond au support de l'image et les feuilles à des extrema locaux, pour favoriser la segmentation des régions spectralement extrémales de l'image telles que les tumeurs dans la TEP.

Dans le contexte clinique, la fusion TEP / TDM est utilisée pour renforcer la perception visuelle des contours des lésions. La TEP possède une très bonne résolution spectrale / fonctionnelle, tandis que la TDM fournit une meilleure résolution spatiale, et donc une description plus précise de la morphologie tumorale. Certains auteurs ont tenté de reproduire ce mode opératoire adapté à la segmentation multimodale TEP / TDM. En fonction de la manière dont ces deux images sont exploitées, les méthodologies peuvent être classées en fonction de l'influence qu'elles ont sur le résultat de la segmentation. Elles peuvent guider l'ensemble du processus pour une co-segmentation complète ou partielle pour à la fin du processus fusionner des segmentations indépendantes. *Bagci et al.* [1] présentent une approche dans laquelle l'algorithme du marcheur aléatoire est appliqué sur la TEP combinée à une ou plusieurs sources additionnelles pouvant être composées de la TDM et / ou l'IRM pour modéliser une segmentation hybride. Chaque image d'entrée est représentée sous forme d'un graphe intra-connecté par voisinage voxel, ces différents graphes étant inter-connectés par correspondance spatiale entre eux pour former un hypergraphe. *Song et al.* [2] ont proposé une méthodologie bimodale basée sur le *graph-cut* pour la segmentation TEP et TDM en tant que problème de minimisation d'un modèle à champ aléatoire de Markov. La structure des graphes multimodaux est similaire à celle présentée par *Bagci et al.* avec des arêtes intra- et inter-nœuds pour la construction d'hypergraphe.

Nous pouvons supposer, sans perte de généralité et sur la base des différentes configurations présentées ci-dessus, qu'une structure de graphe offre une polyvalence, qu'il s'agisse des relations modélisées ou des relations d'intensité pondérée. De plus, la multimodalité accroît la richesse des données à exploiter dans la structure et la caractérisation des graphes. Ces deux concepts, modèle graphique et multimodalité, sont choisis comme base des méthodologies proposées permettant d'améliorer la précision de la segmentation.

# Chapter 1

# Context

## Contents



## 1.1 Medical context

Despite significant clinical advances in research and technological device development, cancer diseases remain a major health problem present worldwide with an estimated 9.5 million deaths in 2018 [3]. In Europe, cancer has received a particular attention because of representing nearly 10% of the world population but 25% of global cancer cases [4], being prostate cancer for men and breast cancer for woman, the highest in incidence and mortality rates, and lung cancer for both sexes in second position. Nevertheless, head and neck cancer spotlight with 26% of world cases. In Europe, there are approximately 140,000 new head an neck cancer cases (an estimated 4% of the cancer incidence) and 63,500 deaths [5], whereas for United States, 3% of incidences (63,000 cases) and 13,000 of morbidity [6] in 2012. These impacting statistics have increased the interest from many related research centers to find respective improvements and solutions in favor of human being health.

Tumor segmentation in medical imaging, specially in PET, has been an important focus of research to provide automated tools for delineating tumor volume with a higher accuracy and reproducibility. PET has been the main segmentation

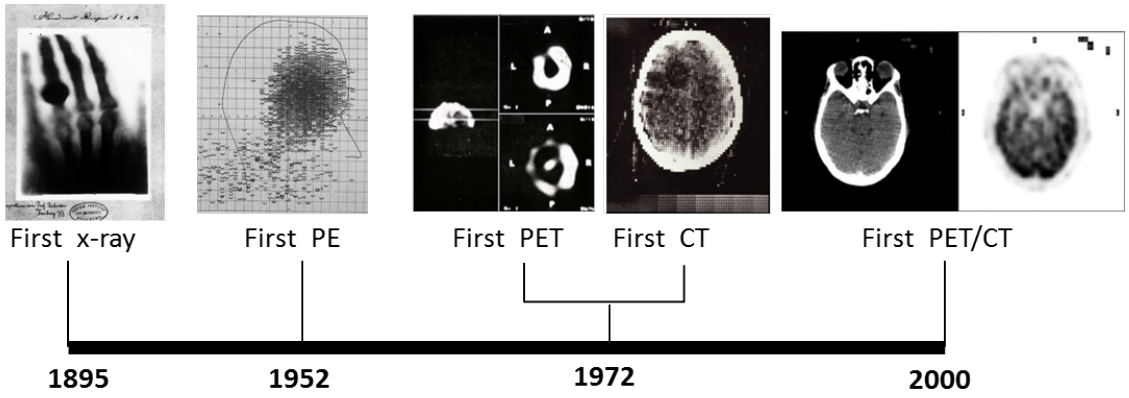


Figure 1.1: PET/CT time line.

reference because of its high contrast between malignant tumors, as a hypermetabolic area, and their background. Recent methodologies incorporate CT data arguing a better performance. The fusion between metabolic and anatomical references allows to evaluate the abnormal activity of tumor cells and in parallel, their anatomic composition and distribution. However, this bimodal framework does not have a common beginning. Their monomodal evolution until their fusion as PET/CT for the establishment of this bimodality is explained in Section 1.1.1, followed by a detailed introduction on PET qualities in Section 1.1.2 to explain its position as a gold standard imaging for several cancer types. The end of this chapter, in Section 1.2, presents the state of the art of approaches presented by other authors which allows us to position our contributions.

### 1.1.1 PET/CT origins

Medical imaging, since the first X-ray in 1895 until the wide range of imaging types nowadays, has revolutionized human body perception (Figure 1.1). Depending on biological data source, medical images can be mainly classified into: anatomical imaging such as X-ray, CT, MRI, ultrasound or functional imaging such as scintigraphy, PET, SPECT, fMRI.

Anatomical imaging refers to the representation of body tissues according to their chemical structure and composition in a molecular level. X-ray was the first anatomical image introduced by *Wilhelm Conrad Röntgen* [7] on an experimentation with discharge tubes. Accidentally, he realized that some invisible rays coming from the tube, passing through its cardboard cover, were reflected on a fluorescent screen one meter away. Then, he discovered its medical application showing an X-ray photograph of his wife's hand. Physically, a X-ray scan is the attenuation of X-rays that are projected through an object in one direction such as coronal or sagittal. In clinic, it allows to contrast dense tissues from soft ones, having as main medical application, the supervision of bone body structures (bone fractures). Computed Tomography or 3D X-ray imaging is the most used anatomical imaging in clinic for 3D human body visualization. CT was introduced in 1972 by *Godfrey Hounsfield* and *Allan Cormack* [8] as an efficient tool to show body tissues by an attenuation map which can be ordered from low to high coefficients as: air, water, muscles, cartilage and bones. Medical ex-

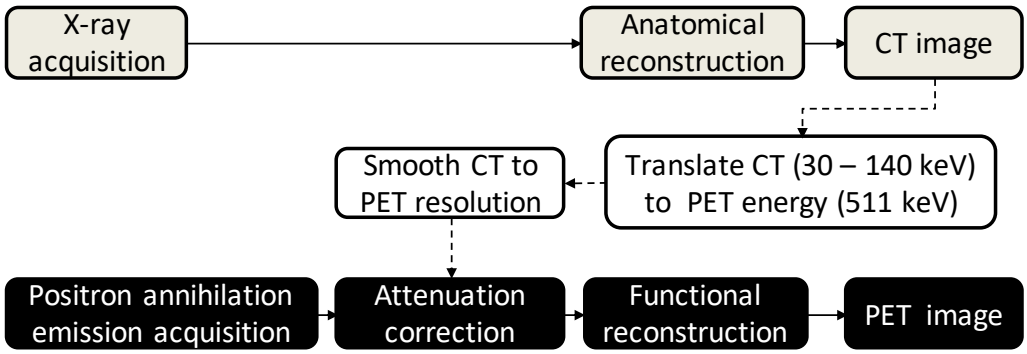


Figure 1.2: Pipeline for PET/CT acquisition.

perts can exploit this information in relation with anatomical disorders such as traumas, musculoskeletal disorders, fractures, infectious diseases and tumors. It can also guide procedures such as surgery, biopsy and radiotherapy or supervise post-treatments patient state. In odontology, dental CT scans are used to analyze maxillofacial area for temporomandibular joint disorders, dental anomalies and/or implant planning. Others domains have exploited CT information too. In Archeology, the first CT scan on a mummy was performed in 1977 by *Derek Harwood-Nash* on the desiccated brain of a boy and a young woman within her cartonnage [9]. For the study of the Ice Man Otzy, probably the most famous recent discovered mummy of 5300 antiquity years, found in the Austrian Alps in 1991, a total of 2190 CT scans were acquired [10, 11].

Functional imaging is the representation of a cell metabolism type on a pictogram. The first Positron Emission (PE) by annihilation radiation imaging applied in medicine was realized in 1951 [12]. It was motivated to study brain activity, mainly to analyze brain tumors. For its part, with the same motivation analyzing brain tumors, Positron Emission Tomography (PET) was presented by *Gordon L. Brownell* with the PC-I in 1972 [13]. The obtained images where promising but with low quality in terms of spatial resolution, and not efficient diffusion and attenuation correction. These drawbacks delayed PET evolution.

The proposal to combine PET with CT was made in the early 1990s by *Townsend* [14], which comprised rotating banks of bismuth germinate (BGO) block detectors whose gaps allowed to incorporate a PET scanner. It was not until 1998 when the first prototype PET/CT scanner with a single-slice spiral CT scanner and a rotating ECAT ART PET scanner, introduced by Beyer [15], became operational and quickly became popular for clinical application. Using CT in a sequential acquisition provides an attenuation coefficient map spatially related with PET volume which can directly help to correct PET attenuation artifacts [16] (Figure 1.2). Secondly, this spatial common framework can be used for PET/CT co-visualization producing a higher anatomical orientation of cellular metabolism.

PET/CT is widely used to evaluate metabolically traceable diseases, such as endocrine, neurological disorders and several kinds of cancer diseases. CT itself is limited on cancer diagnosis. Cancerous cells often do not have a high contrast with respect to their surrounding tissues with similar composition. Indeed, an abnormal tissue could be just detected looking at anatomical tracks, in terms of

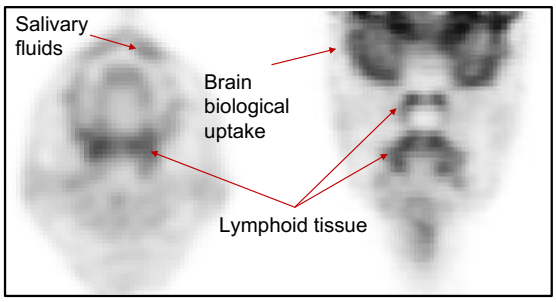


Figure 1.3: Normal high uptakes on an axial–coronal PET.

position and symmetric coherence. PET provides a more sensitive data of cancer cells. It uses radioactive products, generally called radiotracers or radiopharmaceuticals, to mark the cellular metabolism. Normal cells uptake this biomarker, but abnormal cells have, respectively, higher uptake. $^{18}F$-Fluordeoxyglucose (FDG) is the most used PET radiotracer in clinical purpose [17]. Glucose uptake is supervised since it is the main source of energy for cells. Cancer cells increase FDG uptake because of their aggressive growing behavior.

However, other facts can determine a high uptake of this radiotracer [18] (Figure 1.3). Brain is an organ which metabolism is glucose dependent. Therefore, very intense radiotracer fixation happens for cerebral cortex and basal ganglia in a normal state, representing the 7% of total uptake of injected dose. Moreover, lymphatic system nodes can display a high uptake since they are station of production and transit of immune cells, thus glucose consumption. Palatine and lingual tonsils are examples of this behavior, where the relationship between size

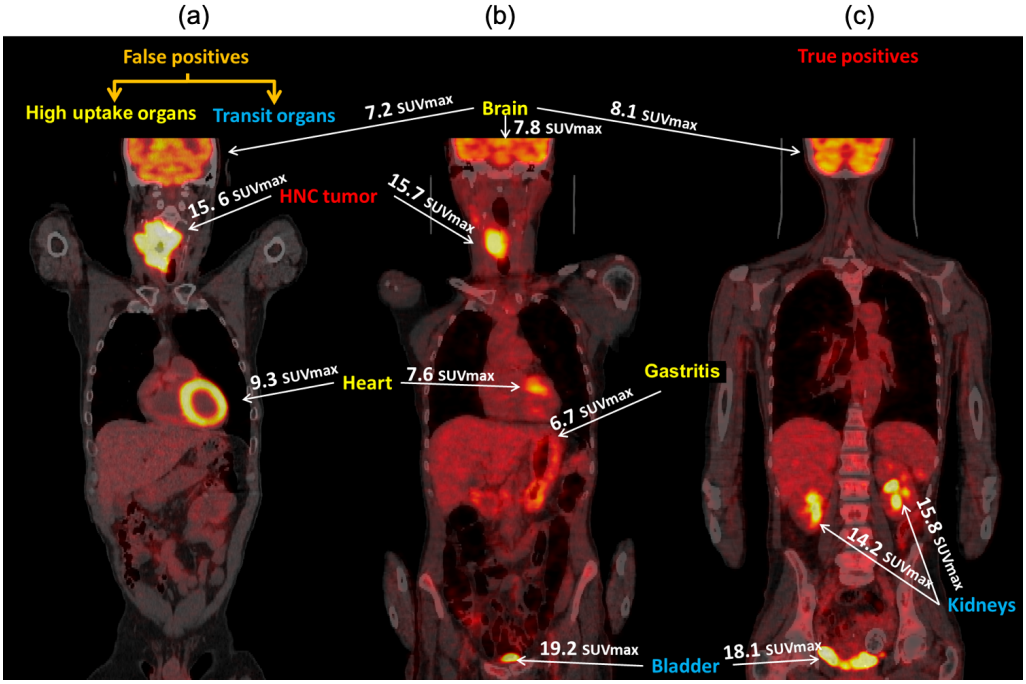


Figure 1.4: Interpretation of SUV on 3 fused PET (heat colormap)/CT (graylevel) images in coronal view where (a) and (b) are head and neck cancer patients and (c) is a healthy subject.

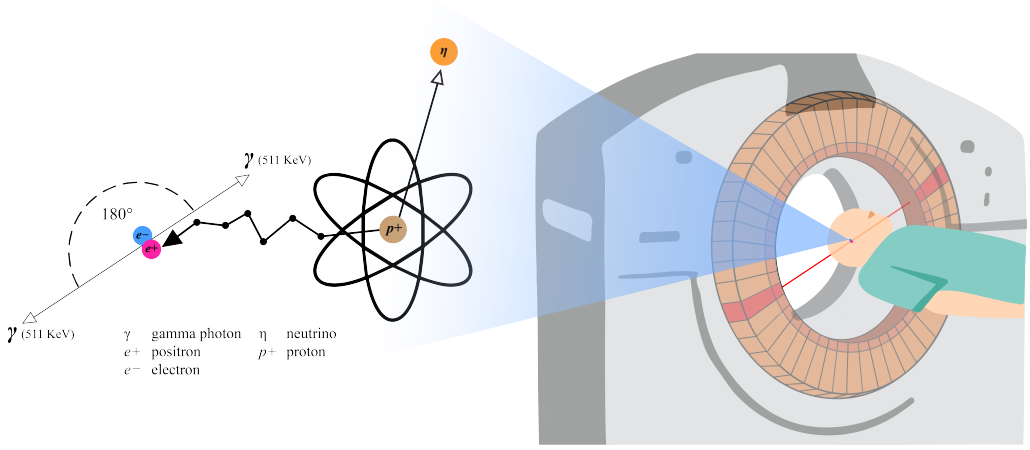


Figure 1.5: Positron/electron photons detection for PET imaging reconstruction.

and uptake can help to identify a normal high uptake. Other non-cancerous high uptake can be displayed on tissues under inflammatory response. Gastrointestinal tract can have an intense biomarker uptake due to gastritis (Figure 1.4.(b)). Moreover, the normal heart, specifically talking about myocardium, can use both glucose and fatty acids as sources of energy. It possesses a metabolic flexibility according to its need and the physiological offer to use adenosine triphosphate suppliers. This condition determines heart FDG uptake that can be high on normal state (Figure 1.4.(a,b)).

Other false positive situations can be detected on the excretion path of liquids from the body. Kidneys such as the blood filtering pair organs and the bladder such as the final urinary elimination deposit, both regularly show a high intensity in PET images (Figure 1.4.(b,c)). Possible tumors in these zones can be totally or partially hidden, complicating their detection. Another drawback on PET interpretation is the partial volume effect (PVE). It is a consequence of limited spatial resolution in reconstructed PET images. It is to say that internal contrast of volumetric area provokes a fuzzy representation. Thus, maximal intensities of signals have the tendency to be underestimated. This a big problem for small tumors and generally for tumor to background transition zones as boundaries.

### 1.1.2 PET fundamentals

Positron Emission Tomography allows for 3D metabolism representation [19], produced by detection of coincidence gamma photons coming from electron–positron annihilation. Depending on evaluating metabolism, a biomarker bounded with a $\beta^+$ decay radionuclide is selected to be injected to the patient. A physical-chemical process, through internal body, is carried out. Biomarker is guided through the network of vessels by blood flow, entering the cells by their respective metabolism or being evacuated from the body. During all this procedure, a nuclear disintegration can take place in the radioactive marker whose unstable nucleus, in order to obtain an equilibrium state, transforms a proton into a neutron, releasing two particles, a small one (such as an electron but with positive charge), called positron; and a neutrino (with no charge). The positron then travels a few millimeters, losing kinetic energy until remaining an energy of

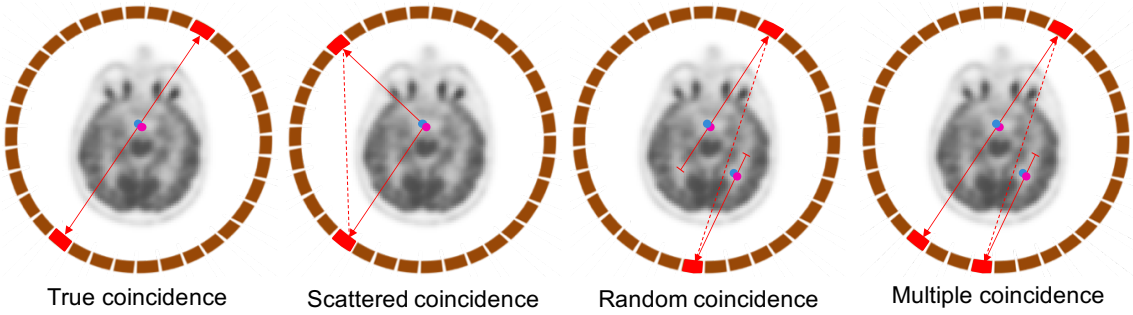


Figure 1.6: Types of coincidence that can be detected in PET.

511 keV to be annihilated with an electron (Figure 1.5).

Thus, two photons of the same energy (511 keV) are released in a random and opposite direction. A ring of sensors around the body captures these photons, detecting and storing their incidence. The relation of two photons coming from the same annihilation event (coincidence event) build a line of response (LOR). The recognition of coincidence events is not perfect due to photon interaction through body tissues. Photons can be deviated or totally attenuated in their path through the body, which provides a high noise expressed on false coincidences or a high reduction of LORs coming from high density zones.

From capture of photons until final PET version, there are some steps for improving image quality before and after reconstruction. Pre-reconstruction treatment is mainly focused on how to reduce false coincidence events since photon capture, followed by the classification and selection of true coincidence events. After reconstruction, lost information from attenuated photons can be estimated and corrected to provide a higher metabolic approximation. All this sequential procedure is detailed below.

### Coincidence problems

Positron annihilation is recognized along a LOR of 180 degrees by sensor rings that are axially positioned representing $z$ slice dimension. This coincidence recognition can be classified in 4 categories: true, scattering, random and multiple coincidence.

(1) A true coincidence is the most suitable case but not the most common since photons are subject to physical factors before being detected (Figure 1.6). Photons can be susceptible to radiation-matter ionizing phenomena in their passing through body mass in their direction to sensors, being Compton diffusion and attenuation the main drawbacks.

(2) They are affected by Compton scattering when an electron deviate gamma photon trajectory. Therefore, it is highly possible to have a wrong LOR assignment for the resulting coincidence event. In PET, more than half of the detected coincidences are scattering coincidences. Therefore, it is essential to correct diffusion. There is not an exact correction, but several approaches have been proposed like the methods based on model exploitation [20].

Moreover, photons can be affected by attenuation by media absorption, thus they do not reach the sensor. If just one photon of an annihilation event couple is totally attenuated, noisy signals can be added to probably generate false LORs.

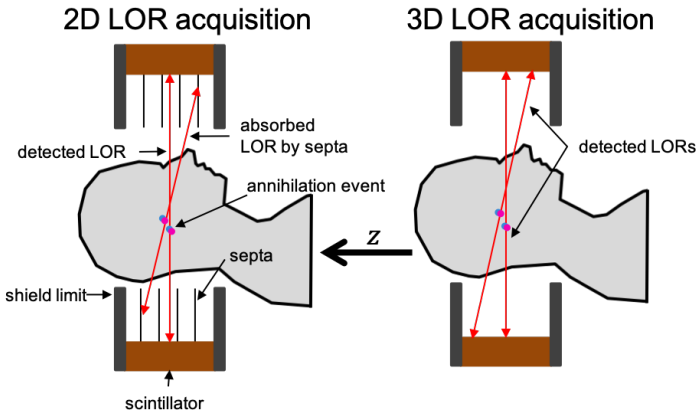


Figure 1.7: 2D and 3D acquisition modes.

(3)Two incident photons reaching the sensor but coming from different annihilation events, within coherent temporal window, can provoke a random coincidence. (4) Moreover, if more than two photons are incidents in different sensors within time coincidence window, multiple possible coincidences are obtained.

### Capture mode of coincidence events

Capture of photons can be done in 2D or 3D mode (Figure 1.7). For 2D mode, thin ring, called septa and generally made of tungsten, axially disposed between cristal rings detectors, are used to block crossing photons, just detecting LORs from a specified imaging plane. For 3D mode, septa are removed and detection inter-rings is allowed. The 2D mode is easier to reconstruct blocking noisy data by septa collimation. However, sensitivity decreases because of blocking also true coincidences. Conversely, for 3D mode, not septa version, there is a larger number of either true coincidence than scattered and random coincidences, increasing sensitivity but getting harder to make the reconstruction due to higher noise.

### Time of flight

Time of flight (TOF) [21] is a technique used to improve the detection of coincidence events. Its raw definition is simply the subtraction between the time coincidence detected ($\Delta t$) of two gamma photons. The nearest photon logically will first reach a detector, and time coincidence difference $d$ allows to approximately calculate the position of annihilation event along the respective LOR. However, this time-spatial conversion depends on time coincidence resolution $t_{res}$ of the captors, which defines an uncertainty $U_{TOF}$ by:

$$U_{TOF} = \frac{t_{res} * c}{2}, \tag{1.1}$$

where $c$ is the universal physical constant of speed of light. Thus, the annihilation event position event $d$ has an uncertainty $\pm U_{TOF}$:

$$d = \frac{\Delta t * c}{2} \pm U_{TOF}. \tag{1.2}$$

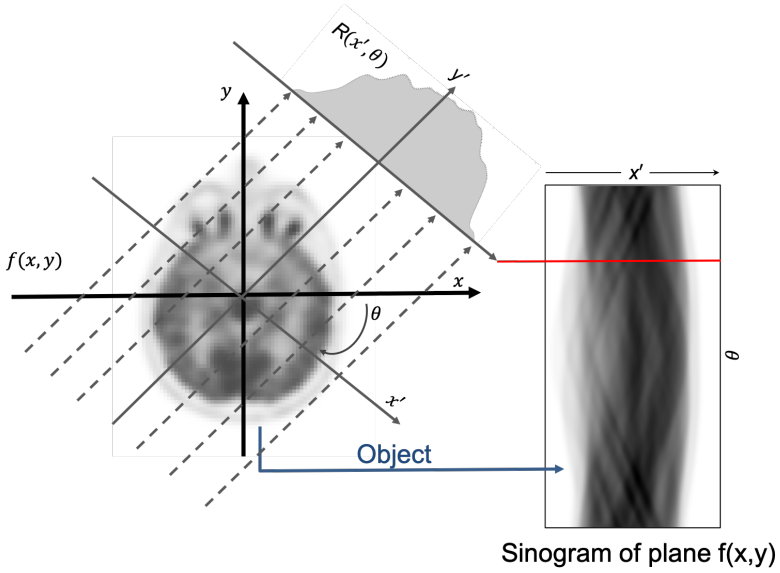


Figure 1.8: Image plane $f(x,y)$ represented by a sinogram from a Radon transform $R(x',\theta)$.

where $\Delta t$ is the arrival time difference on the first and second detectors. Ideally, if we know whole $d$ original annihilation positions, PET reconstruction can be directly applied without need of using iterative algorithms. Indeed, image 3D matrix can be filled by $d(x,y,z)$ positions. However, $U_{TOF}$ spatial uncertainty is still considerable nowadays (around 1.5-9.0 cm) due to 90-600 picoseconds time coincidence resolution detectors [22].

**PET reconstruction**

PET follows the tomographic reconstruction basis exposed in the Radon transform. An object can be represented by the set of its radial projections $R(x',\theta)$ of constant length planes $x'$ with a degree $\theta \in \mathbb{R}[0,180]$ called sinogram (Figure 1.8).

Applying the inverse process, generally called backprojection the object can be reconstructed as follows:

$$b(x,y) = \int_0^{\pi} R(x',\theta)\,|_{x'=xcos\theta+ysin\theta}\, d\theta, \tag{1.3}$$

where the reconstructed image $b(x,y)$ is blurred by image slowly rolls off, thus composed of original $f(x,y)$ plus noise:

$$b(x,y) = f(x,y) \times \frac{1}{\sqrt{x^2+y^2}}. \tag{1.4}$$

Based on this concept, different analytic and iterative algorithms have been proposed to carry out the discrete reconstruction of the image through the acquired sinograms and to solve the blurring problem.

Filtered backprojection [23] is a well-established analytic method due to low computational cost and high performance. Sinogram is firstly changed into spectral domain by Fourier transform, usually by Fast Fourier Transform (FFT), to

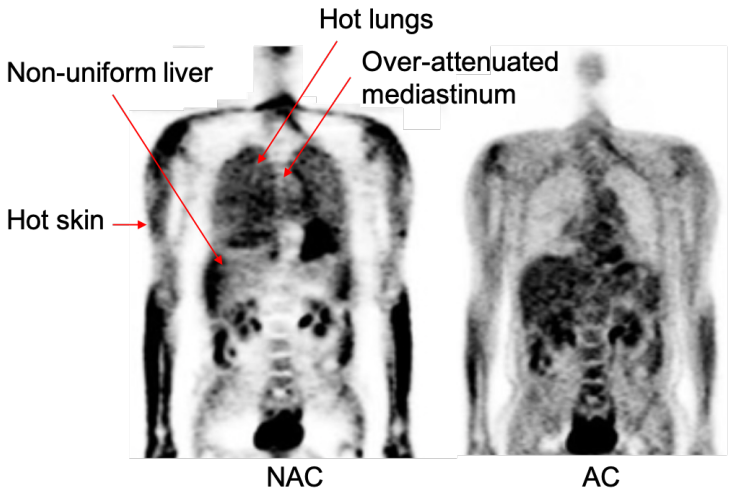


Figure 1.9: Biological attenuation problems in non-attenuation correction (NAC) PET and its improved version after attenuation correction (AC).

apply a ramp filter for deblurring. The filtered sinogram, recovered from inverse FFT, is backprojected for reconstruction.

Modern iterative algorithms with high computing, have been proposed to improve the quality of reconstructed image, where Ordered-Subsets Expectation Maximization (OSEM) is the most popular [24].

#### Attenuation correction

The most used type of different level PET corrections is the Attenuation Correction. As explained above, photons, according to relative annihilation path, can be completely absorbed or attenuated. In other words, photons have a heterogeneous attenuation depending on tissue they crossed. Therefore, those ones which go through dense tissue are more likely to be attenuated than soft tissues. For example, superior mediastinum often presents a higher attenuation being denser than lungs, and caged in the front by the sternum and backside by first thoracic vertebrae (Figure 1.9). By contrast, lungs exhibit an overestimated uptake, as it is also the case for skin, the outer body layer which has lower attenuation drawbacks. Moreover, both lower and higher contrast issues are frequent in certain areas on non-corrected images, for example in the liver, due to different attenuation despite similar uptake.

In essence, attenuation problem causes a non-coherent radiotracer uptake quantification. To correct this problem, attenuation coefficients through the body mass must be defined for all LORs. For PET/CT scanners, the coupled CT is used for PET attenuation correction due to its meaningful advantages. CT acquired in a short time difference provides accurate information to co-reproduce patient position for PET-CT matrices and be adapted *in vivo* to local conditions (e.g. artifacts, physiological abnormalities), allowing correction for attenuation. However, CT attenuation coefficients have to be modified according to PET energy photons because of different interaction with matter. For CT with 30–140 keV photons, there is a combination of Compton scattering and photoelectric absorption while for PET monochromatic 511 keV photons are mainly affected by Compton scat-

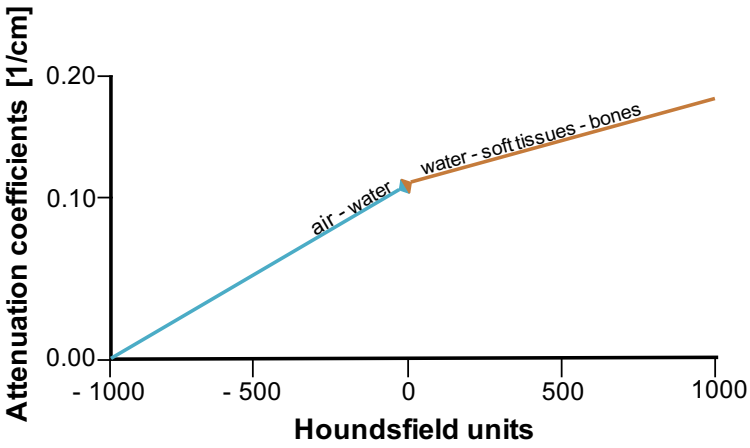


Figure 1.10: Bilinear scaling to convert CT to PET attenuation coefficients.

tering. Attenuation coefficients are highly similar for non-bone tissues where Compton effect is highly relevant, while bone has a higher photoelectric effect because of calcium level. Thus, bilinear scaling method is widely used as Attenuation Correction treatment (Figure 1.10).

Although spatial correspondence between PET and CT is attempted from co-acquisition, it is not often reached. CT is acquired in a short time, less than 1 minute approximately, thus thorax is imaged in one particular single position. However, for PET which acquisition takes around 15 minutes, thorax image is related to a mean of multiple positions, and there is a great generation of respiratory artifacts in the thoracic area. Therefore, a considerable false partial correction can be obtained applying CT attenuation correction [25]. In Figure 1.11, we can see an example of attenuation correction common error. In (b), PET and CT

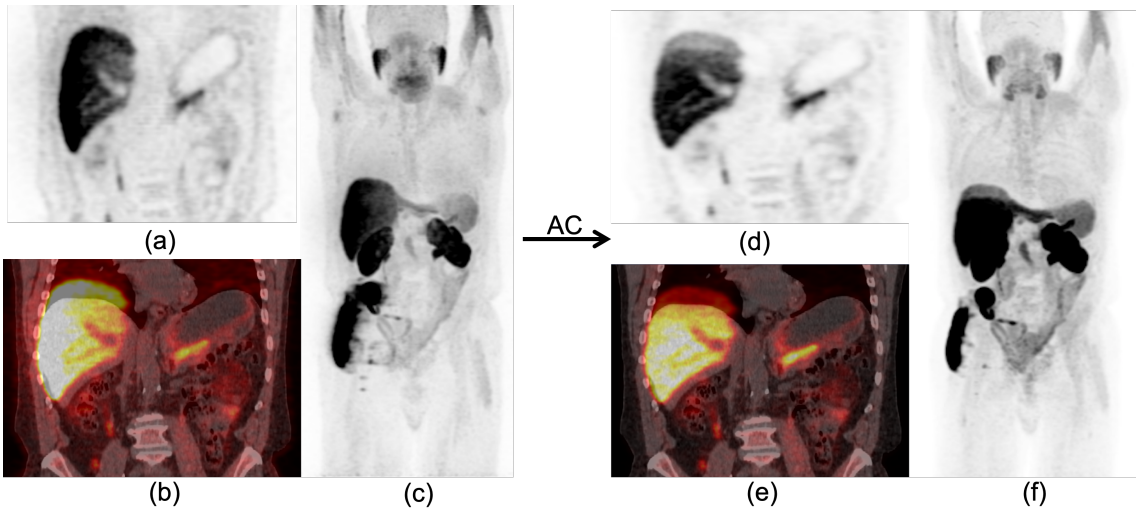


Figure 1.11: Partial error in PET attenuation correction by CT clearly shown on liver on a PET/CT recovered from the Godinot Institute database. (a) Non-corrected PET, (b) PET/CT and (c) maximum intensity coronal projection of PET and (d,e,f) their respective attenuation correction (AC) versions.

| Tag | Keyword | Value representation |
|---|---|---|
| (0008,0031) | SeriesTime | int vector (Time = [days hours seconds]) |
| (0010,1030) | PatientWeight | Decimal String (Kilograms) |
| (0018,1074) | RadionuclideTotalDose | Decimal String (Bequerels) |
| (0018,1072) | RadiopharmaceuticalStartTime | int vector (Time = [days hours seconds]) |
| (0018,1075) | RadionuclideHalfLife | Decimal String (seconds) |
| (0028,1053) | RescaleSlope | Decimal String |

Table 1.1: DICOM metadata for SUV computing.

overlapped livers show a spatial offset as respiratory artifact. This error is projected to (d) by assigning partially incorrect attenuation coefficients in correction process. Maximal intensity projections (c,f) allow to confirm the error propagation through slices. This respiratory motion problem can be solved by applying a thoracic PET/CT co-registration [26].

**Standard Uptake Value**

Standard Uptake Value (SUV) [27] has been proposed as a metric to determine the difference between normal and abnormal uptake levels. It involves a normalization of PET intensities according to the distribution of radioactivity of biomarker injected through the body. It procures a reproducible metric defined by:

$$\mathrm{SUV} = \frac{C_{img}}{C_{inj}}, \tag{1.5}$$

where $C_{img}$ is the local concentration at image voxel measured in kBq/mL and $C_{inj}$ is the ratio between the injected dose (kBq) and patient weight ($g$).

SUV can be easily computed in DICOM images from their data and metadata. Indeed, DICOM data can be SUV transformed recovering specific parameters in the file format metadata. This information can be found globally equal ($gbl$) in all axial slice files or locally ($lcl$) from slice specifications. 2D data slice, represented by gray levels, can then be converted to concentration $C_{img}$ by:

$$C_{img} = \mathrm{GrayLevel} * \mathrm{RescaleSlope}_{lcl}. \tag{1.6}$$

For $C_{inj}$, parameters can be fully obtained on metadata and thus computed by:

$$C_{inj} = \mathrm{RadionuclideTotalDose}_{gbl} * \mathrm{DecayFactor}_{lcl} / \mathrm{PatientWeight}_{gbl}, \tag{1.7}$$

where PatientWeight dimension, using 1g = 1mL equivalence from water density, adjust the SUV as an adimensional metric. DecayFactor allows to relatively adapt the injected dose decay at time each image series is acquired:

$$\mathrm{DecayFactor} = 2^{-(\mathrm{SeriesTime}_{lcl} - \mathrm{StartTime}_{gbl}) / \mathrm{RadionuclideHalfLife}_{gbl}}, \tag{1.8}$$

where Series Time refers to slice acquisition start, Start Time to exam acquisition start and Radiopharmaceutical Start Time to the time radiomarker was injected. Main metadata parameters mentioned above are detailed in Table 1.1.

| Radionuclide | Half-life (m) | Decay type | Energy (MeV) | Decay isotope |
|---|---|---|---|---|
| $^{11}Carbon$ | 20.33 | $\beta^+$(99%), EC (0.23%) | 0.96 | $^{11}Boron$ |
| $^{13}Nitrogen$ | 9.97 | $\beta^+$(99%) | 1.20 | $^{13}Carbon$ |
| $^{15}Oxygen$ | 2.07 | $\beta^+$(100%) | 1.72 | $^{15}Nitrogen$ |
| $^{18}Fluor$ | 109.70 | $\beta^+$(96.86%), EC (3.14%) | 0.63 | $^{18}Oxygen$ |
| $^{64}Copper$ | 720.86 | $\beta^+$(17%), $\beta^-$(39%), EC (43%), $\gamma$ (0.47%) | 1.72 | $^{64}Nickel$ ($\beta^+$, EC) and $^{64}Zinc$ ($\beta^-$) |
| $^{68}Gallium$ | 68.25 | $\beta^+$(100%) | 1.92 | $^{68}Zinc$ |
| $^{82}Rubidium$ | 1.27 | $\beta^+$(99%) | 3.15 | $^{82}Kripton$ |

Table 1.2: Positron emitter radionuclides and their main features.

In relation with SUV interpretation, theoretically, if the biomarker was spread homogeneously in the body, SUV must be equal to 1 everywhere. However, biomarker metabolism of body cells is different depending of particular need they have on biomarker: lower metabolism for fat cells, higher metabolism for the active brain regions and carcinogenic cells. A SUV of 2.5, under normality exceptions, has been defined as a common threshold for abnormal metabolism [27]. A variation of SUV is based on lean body mass concept $SUV_{LBM}$ defined as the total body weight minus fat weight due to fat sensitivity for a lower uptake, thus disturbing total uptake distribution. Indeed, other factors as temperature, age, sex or diabetes can also disturb SUV value [28].

### Radionuclides

A radionuclide is the radioactive core of a biomarker. Chemically, it is an atom isotope whose nucleus is unstable, thus emitting particles/energy. Its instability can be classified according to disintegration type, half-life and decayed isotope. The type of released particles define the type of disintegration such as $\alpha$ for $\alpha$ particles (2 neutrons + 2 protons, as an Helio atom), $\beta^-$ for electrons, $\beta^+$ for positrons, $\gamma$ for $\gamma$ photons or electron capture (EC) for proton neutralization by electron absorption, for example. The half-life of a radioisotope, particular for each nuclide, is a statistical measure of the time required for half radioactive nucleus disintegration. In other words, the probable time needed for the half set nucleus decay. This decay phenomenon can be modeled according to the universal law of radioactive decay:

$$N(t) = N_0 2^{-t/t_{1/2}}, \tag{1.9}$$

where $N(t)$ refers to remaining radioactive nuclei after a time $t$ with an initial state $N_0$ and a half-life $t_{1/2}$. The decayed nucleus represents a new type of chemical element with a new nuclear configuration, which is classified according to the type of particle that was disintegrated (Table 1.2).

For PET, radionuclides are characterized to be positron emitters ($\beta^+$). Fluorine 18 is the most used radioisotope in PET oncology due to its intrinsically suitable features: a favorable half-life to be produced, transported, applied on the patient, captured during the exam and rapidly evacuated from patient. Moreover, its initial positron energy (630 keV) causes a short path until 511 keV annihilation, thus providing a low spatial uncertainty.

| Biomarkers | Metabolical process | Clinic usage |
|---|---|---|
| $^{18}F$-fluoro-deoxy-glucose (FDG) | Glucose metabolism | General |
| $^{18}F$-FNa | Osteogenesis | Bone metastasis |
| $^{18}F$-fluoro-dihydroxy-phenylalanine (F-DOPA) | Protein synthesis | Pheochromocytoma, paraganglioma, neuroblastoma, neuroendocrine tumors (midgut), glioma |
| $^{18}F$-fluoro-choline (FCH) | Membrane lipid synthesis | Prostate cancer |
| $^{68}Ga$-PSMA | Prostate specific receptors | Prostate cancer |
| $^{68}Ga$-DOTA-peptides | Somatostatine receptors | Neuroendocrine tumors |
| $^{18}F$-fluoro-thymidine (FLT) | DNA replication | General research |
| $^{18}F$-FES | Estrogen receptors | Breast cancer research |
| $^{18}F$-fluoro-galacto-RDG | Angiogenesis | General research |
| $^{18}F$-fluoro-misonidazole (FMISO) | Hypoxia | General research |
| $^{18}F$-fluoro-annexin V | Apoptosis | General research |

Table 1.3: Common biomarkers for clinical and research use.

In nuclear medicine there are three main methods to produce radioisotopes: nuclear fission, neutron reactor and bombardment by charged particles. The first 5 isotopes mentioned in Table 1.2 can be obtained from charged particle bombardment in cyclotrons. $^{64}Cu$ can be also produced from $^{64}Zn$ by a nuclear reactor where atoms are bombarded with medium to high energy neutrons. $^{68}Ga$ and $^{82}Rb$ are directly obtained by disintegration of $^{68}Ge$ and $^{82}Sr$ generators.

### Biomarkers

A radiotracer, radiopharmaceutical or biomarker is a radioactive substance used to mark a physiological activity and obtained by bounding a radioactive isotope with a molecule of a biological substance such as glucose in the case of $^{18}F$-FDG. In oncology, the radiotracers allow to highlight abnormal metabolic activities, which is a regular feature for carcinogenic cells. Cells need energy to fulfill their functions. They increase the amount of entering glucose and of enzymes to metabolize it. The FDG enters inside the cell and is phosphorylated by hexokinase in fluorodeoxyglucose-6-phosphate. However, the $^{18}F$ bond replacing hydroxide molecule prevents biomarker to be degraded by the body to complete glucose metabolism. This state prevents it from leaving the cell and, thus radioactivity is blocked inside which favors its suited place caption. $^{18}F$ decays to $^{18}O$ to form a molecule of glucose-6-phosphate, reaching nuclear stability (non-radioactive) without consequences for metabolic functions.

Fluorodeoxyglucose ($^{18}F$-FDG) is the most used radiotracer in clinic [29] for the diagnosis of several types of cancers such as lung cancer, breast cancer, lymphoma, cervicofacial cancer, genitourinary cancer, melanoma or sarcoma. However, there are cases where FDG is not convenient. The choice of a particular radiotracer depends on the anatomical region and cell type of the studied pathology. Brain cancers, due to high glucose uptake on brain activity, have a low contrast to be easily analyzed by FDG. Also for prostate tumors, FDG uptake is more inconstant due to a cellular efflux of this radiotracer. However, F-DOPA based on amino acid tracer can provide better results for brain tumors by higher decreasing on normal neighborhood uptake rate [30].

In Table 1.3, a list of radiotracers are presented. Their use is specified ac-

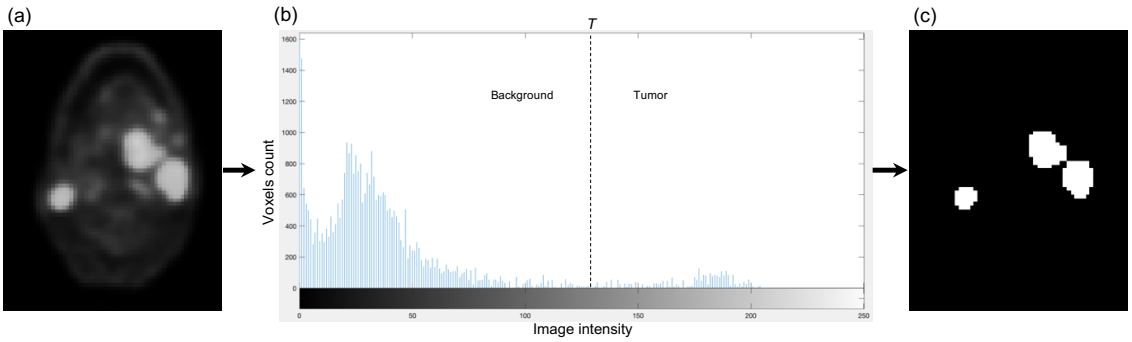


Figure 1.12: Thresholding by pixels higher than $T$, from (a) PET axial slice at maxilar level, trough (b) histogram for (c) tumor segmentation.

cording to a type of pathology. General term refers to a big range of pathologies studied with that have low exceptions.

## 1.2 State of the art for PET segmentation

The analysis of tumors is essential for staging and treatment planning for patients with cancerous diseases. PET/CT provides high quality data to be explored *in vivo* by medical experts. Fused bimodal visualization increases tumor tracks for handmade computer assisted delineation. This interactive and intuitive manual segmentation is widely applied to define regions of interest (composite zones, organs, tumors) for object feature extraction and radiation therapy planning. However, this manual implementation is time-consuming. Moreover, since operator judgment intervenes in the process, there is a level of reproducibility lack coming from the same operator (intra-error) or from different operators (inter-error), to define which voxels belong to foreground and which ones to background.

Several methodologies have been developed [31] to handle these constraints, to optimally automate this operation, looking for reproducible and accurate results reducing time-consumption; and overcome intrinsic image limitations.

Monomodal PET segmentation approaches have been firstly and highly developed taking advantage of tumor high contrast. These methodologies can be classified in relation with their relative domain on: threshold-based, contour-based, stochastic and graph-based approaches. Recently, multimodal approaches have been proposed, increasing segmentation accuracy by exploiting different data sources, mainly functional and anatomical sources. These approaches are presented below, describing their main advantages and drawbacks.

### 1.2.1 Threshold-based approaches

Thresholding segmentation is based on the selection of a value in the histogram of intensities, whose greater voxels are considered as foreground and the remaining ones as a background, converting a gray-level image into a binary one (Figure 1.12). In clinical practice, these approaches are popular for PET tumor segmentation due to their simplicity to be used under favorable conditions to retrieve satisfying results. Indeed, PET has as main characteristic: a high tumor to back-

ground contrast. Thus, high uptake regions correspond to high gray-level intensities, which can be sharply separated from their background by a threshold value.

However, the main difficulty lies in the choice of the threshold. This selection can be intuitively performed by empirical decision according to the type and context of cancer, or automated by phantom calibration or iterative algorithms. The different techniques can be classified on fixed, adaptive and iterative approaches.

### Fixed thresholding

A fixed threshold is a direct-cut threshold to be applied on the image. This threshold value can be defined, intuitively, observing the histogram of intensities or by a pragmatic decision. In clinic, specific experimentations have been held to answer the question of "the best threshold" for several cancer types, to separate malignant lesions from background. These studies have predefined a standard reference threshold catalog. A percentage threshold based $T(\%)$ of the maximum intensity $I_{max}$ is commonly used, which standard interval is between 40-50% [32, 33]. A 2.5 SUV threshold has been applied too as the beginning threshold to consider a high uptake [34]. However, other authors have suggested high contrasting thresholds such as 78% [35]. Indeed, these referenced fixed values are not always efficient, thus are forced to be considerably modified for different PET images and tumor cases. This behavior can be considered normal from the idea of defining stable parameters (fixed thresholds) for a dynamic framework (tumors in PET images) that is multifactor dependent (body physiology, image properties, noise, artifacts).

### Adaptive and iterative thresholding

In an attempt to adapt threshold selection according to input conditions, several analytical and iterative approaches have been proposed. The main idea is to take into account specific characteristics of each image in order to define an optimal threshold $T(\%)$ to segment the object of interest.

For analytical adaptive threshold, prior parameters are injected in an equation to compute $T(\%)$. This equation can be calibrated by a phantom-based ground-truth by using artificial or physical phantom experimentation. The discussion between authors is focused on the definition of parameters which are relevant for tumor segmentation.

*Shaefer et al.* [36] argued that threshold depends on background (BG), $\mathrm{SUV}_{max}$, $\mathrm{SUV}_{70}$ input parameters:

$$\frac{a * m * \mathrm{SUV}_{70} + b * BG}{\mathrm{SUV}_{max}}, \tag{1.10}$$

where $a$, $b$ and $m$ are computed by linear regression, $\mathrm{SUV}_{70}$ is the 70% of $\mathrm{SUV}_{max}$ and $BG$ is obtained by SUV average of selected background regions. Experimentation was applied on a cylindrical polymethylmethacrylate phantom enclosing glass spheres and incremental from 2.5 to 33 signal to background ratios. Spherical volumes have $-0.4$ to $+0.7$ mm error in radius, while PET gross tumor volume has $-0.7$ to $+1.2$ error compared with CT ground-truth.

*Nestle et al.* [37] defined a relationship incorporating an addition between object and background data:

$$T(\%) = a * SUV_{70mean} + (b - a) * \overline{BG}, \tag{1.11}$$

where $a$ and $b$ are obtained by phantom calibration, and $SUV_{70mean}$ was selected to ensure measure inside the tumor. $SUV_{mean}$ of 20-60 $cm^3$ regions of interest localized in the mediastinum, was defined as $\overline{BG}$. Experimentation on Non–Small Cell Lung Cancer (NSCLC) patients showed a higher accuracy of this performing than $SUV_{40max}$ and 2.5 SUV fixed thresholds, based on a manual delineation ground-truth.

*Erdi et al.* [38] expose an inverse exponential relation between volume $V$ and optimal threshold $T$, guided by extracted signal-to-background coefficients:

$$T(\%) = A * e^{-C*V}, \tag{1.12}$$

where $A$ and $C$ are regression parameters from estimated volume $V$ on CT image. These latter are obtained by calibration on an elliptical Jaszczak phantom with 0.4 to 5.5 $cm^3$ spheres that are filled with S/B ratios of 7.4, 5.5, 3.1, and 2.8. Experimentation was applied on lung lesions with a mean difference of 8.4% between these lesions.

To improve the performance of adaptive thresholding, iterative approaches have been proposed, converging on the optimal threshold. Calibration for these methods is quite similar, where the parameters $a$ and $b$, depending on the ratio of signal-to-background, are obtained by linear regression from the calibration curves.

*Jentzen et al.* [39] introduced an optimized version to omit tumor volume input $V$, generally obtained by CT delineation in clinical application. In their approach, at each iteration $i$, $V(T_{i-1})$ is injected to compute $T_i$ until $V(T_i)$ remains invariant between successive iterations. The threshold $T$ can be here defined as follows:

$$T(\%) = a + \frac{b}{V} \tag{1.13}$$

On phantom experimentation, a 10% overestimation in spheres larger than $a$ $cm^3$ is obtained, while for smaller ones this deviation increases. PET images from patients with different varieties of cancers (lung, head and neck, and gastroinstestinal) were used for clinic experimentation. Lesions between 0.8-7.5 $cm^3$ had a 9% average deviation while 15% for larger volumes.

*Daisne et al.* [40] argued that optimal $T$ dependency was solely based on background data:

$$T(\%) = a + \frac{b}{S/B}, \tag{1.14}$$

where $B$ is the average of one or more regions of interest for background near to the tumor and $S$ is the average of voxels surrounding $SUV_{max}$ voxel. Artificial experimentation on spherical lucite phantom showed an overestimation for the smallest sphere (1 $cm^3$) until 67% and largest sphere (17 $cm^3$) got a 13% underestimation.

Typically, the volume of the larger spheres was slightly underestimated whereas the smaller spheres were overestimated.

Inversely, *Black et al.* [41] argued that there was a strong linear relationship between $\text{SUV}_{mean}$ of $V_i$ and the optimal threshold, ignoring background information:

$$T(\%) = a + b * \text{SUV}_{mean}. \tag{1.15}$$

Phantom and NSCLC experimentation with this method were compared with 42% threshold, obtaining a 21% and 67% difference respectively. The result of this threshold model shows a strong linear relation of $T$ with $\text{SUV}_{mean}$, having a short difference with ground-truth.

### 1.2.2 Contour-based approaches

Contour-based segmentation methods aim to identify the discontinuous areas of an image that correspond to boundaries. For PET, this task is not obvious to achieve because of fuzzy boundaries due to low spatial resolution, resulting in blurring objects/regions in the images. However, some boundary-based approaches have been proposed as well, using pre-processing techniques for image deblurring or for directly taking into account uncertainty. They can be grouped in two categories: deformable models and gradient-based methods.

#### Deformable models

Deformable models concept consist of the deformation of an initial model towards a suitable stable one. Active contours [42] are popular in this field. They are literally based on a dynamic object contour, where an initial boundary is going to evolve, having anisotropic movements, analogically as a snake locomotion, until reaching convergence of an energy function. This energy is composed of two antagonistic terms: internal and external energies. These parameters allow to push and pull the boundary of the object of interest through its neighborhood. The internal energy allows to handle the smoothness of the contour, while the external movement is controlled by the features of interest. The classical methods are highly dependent of initial boundary and sensible to topological variations. A geometric active contours were introduced by [43] where evolving contours are defined by level-set layers. These methodologies presented above in the literature have been adopted to build algorithms for PET segmentation. Geometric version appears for NSCLC tumor delineation [44], applying a previous denoising step to reduce noise faced to gradient susceptibility. *Li et al.* [45] defined the first contour by region growing, followed by level-set evolution.

#### Gradient-based methods

A gradient or intensity difference on an image is focused on defining boundaries, generally characterized by a sudden intensity change.

In PET, gradient-based methods have been applied with the challenge of noisy boundaries and fuzzy objects due to low resolution. Indeed, gradient-based methods are highly sensitive to noise which can be amplified. Despite of this drawback, some methodologies have been proposed. *Geets et al.* [46] applied the watershed transform, clustering the regions defined inside the boundaries. In phan-

tom experimentations, a 10-20% understimation was obtained, while in clinical experimentation for NSCLC, a higher accuracy is shown compared with threshold methods [47].

### 1.2.3 Stochastic approaches

Stochastic methodologies use statistical metrics for segmentation. Several methodologies are presented below with their respective particularities.

#### Fuzzy locally adaptive Bayesian

Based on a Bayesian framework, fuzzy locally adaptive Bayesian (FLAB) approach is an unsupervised method of Gaussian mixtures of an object [48]. The segmentation problem is here evaluated estimating the hidden $X$ from the noisy observation $Y$, where $X$ is the segmentation map for $C$ classes set on an image space $Y$. This $X$ and $Y$ relationship is modeled as a joint distribution $P(X,Y)$ by Bayes formula:

$$P(X|Y) = \frac{P(X,Y)}{P(Y)} = \frac{P(Y|X)P(X)}{P(Y)}. \tag{1.16}$$

The fuzzy concept is defined by the fuzzy belonging for a class in fuzzy level from 0 to 1. For PET segmentation, fuzzy labels are assigned for high uptake and background classes. FLAB seems to be robust for homogeneous objects as spheres on the phantom but for heterogeneous objects, as tumors, there is lower accuracy [48]. This heterogeneity problem is proposed to be dealt with by defining a third class [49] for heterogeneous objects. On simulated tumors, experiments show a $9\% \pm 8\%$ mean error compared with a ground-truth, whereas 2 class FLAB led to errors of $15\% \pm 11\%$.

#### Classification and clustering

Classification and clustering approaches aim to label each pixel/voxel of the image by assigning it to a class. The classification process is divided into three stages: definition of characteristic parameters of each pixel/voxel (intensity for example); definition of the characteristic parameters of each class; definition of decision rules for determining pixel/voxel membership in a class. Two families of classification approaches exist according to the second stage (label definition): supervised approaches and unsupervised approaches (commonly called clustering). In supervised approaches, label characteristics are known. It is a kind of an arbitrary classification by a surveyed database.

In the case of unsupervised approaches, there is no *a priori* information to carry out the classification algorithm. Therefore, learning step is done automatically by clusters. K-means, fuzzy-c-means and support vector machine (SVM) are some examples of clustering approaches.

K-means [50] is the most widely used method among unsupervised probabilistic methods because of its simplicity. It was introduced in 1977 for image segmentation [51]. The algorithm is initialized with respect to a number K of initial centers which can be manually or automatically defined. The hard membership classes $S^i$ (monoclass assignment) are then iteratively defined including

the pixels/voxels which have the lowest distance to their $c_i(t)$ barycenter in a time step $t$. This fact provokes an evolution of the barycenter $c_i(t+1)$ placement in a subsequent time step $t+1$ which in turn launches a new pixel/voxel assignment $S_i(t+1)$. This iterative cycle is stopped when barycenters reach a stable state. The process can be explained by minimizing:

$$\arg\min_{S} \sum_{i=1}^{k} \sum_{x_j \in S_i} \|x_j - c_i\|^2, \tag{1.17}$$

where $\|x_j - c_i\|$ is the distance between $x_j$ and $c$ is the barycenters of class k. The parameter $k$ corresponds to the number of classes and $j$ to the iterator of the respective number of points in subset $S_i$.

Fuzzy-c-means, introduced by [52] and improved by [53], are considered as a soft K-means version. The difference lies in the integration of the principles of fuzzy logic into the membership function defined as follows:

$$w_{ij} = \frac{1}{\sum_{k=1}^{c} \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}, \tag{1.18}$$

where $w_{ij}$ is the set of membership coefficients of a pixel/voxel $x_i$ for each class $c_j$. Thus, pixel/voxel can belong to more than one class according to a membership degree from 0 to 1. Centroids of class $c_k$ are computed considering $w_{ij}$ for each pixel/voxel $x$ by:

$$c_k = \frac{\sum_x w_k(x)^m x}{\sum_x w_k(x)^m}, \tag{1.19}$$

where $m$ is a scalar parameter, called the fuzzifier $m \in \mathbb{R}$ with $m \geq 1$, to define fuzziness level. The solution to this algorithm is defined by:

$$\arg\min_{C} \sum_{i=1}^{n} \sum_{j=1}^{c} w_{ij}^m \|x_j - c_i\|^2, \tag{1.20}$$

where c is the number of classes thus the number of barycenters $c \in C$.

On PET tumor segmentation, *Dewalle-Vignion et al.* [54] presented an approach partially using fuzzy-c-means. This approach merges three different methods: Maximum Intensity Projection (MIP), fuzzy-c-means and possibility theory. Firstly, MIP is applied to increase the image contrast and improve the definition of structures. From an interactive selected volume of interest, using MIP basis, passing a ray through an object, preserving the highest value of crossed voxels, three axis plans corresponding to axial, sagittal and coronal directions, are extracted. A classification step by a fuzzy-c-means is then applied at each map to segment tumor and background classes. These fuzzy memebership maps are used, based on fuzzy theory method proposed in [55] for a retroprojection reconstruction, to build 3 membership volumes. Finally, three fuzzy volumes are fused, according to algorithm explained in [56], to obtain an overall volume. Results show a 0.8 Jaccard index with 40% thresholding.

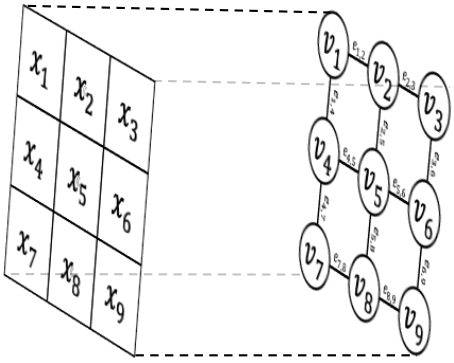


Figure 1.13: Building a connected graph $G(V,E)$ from an image $I$, where pixels $x_i$ are considered as nodes $v_i$ and their 8 neighbors are linked by edges $e_{ij}$.

### 1.2.4 Graph-based approaches

Graph-based approaches generally consist of projecting the image matrix of pixels/voxels to a graph space with nodes and edges, providing a more flexible model to handle gray level intensities and their interaction.

The simplest kind of image graph is to directly consider a pixel/voxel such as a node $v \in V$ which neighbor nodes are connected by edges $e \in E$ for a graph $G(V,E)$ (Figure 1.13). In random walker approaches, this type of graph has been exploited, weighting edges to find the shortest path between marked voxels and unmarked ones for labeling. *Stefano et al.* [57] compared random walker with region growing implementation on PET for head and neck tumors, obtaining a respectively 0.96 versus a 0.91 Dice Similarity Coefficient (DSC) average. *Soufi et al.* [58] applied a random walker approach combined with fuzzy logic on PET images for lung, liver and abdominal tumor segmentation. Results show an improvement until 21% compared to classical random walker. Equally for this image graph concept, graph cut approaches based on min-cut/max-flow paradigm have been proposed. *Ballagan et al.* [59] used an energy based on SUV cost function and monotonic downhill SUV feature. Compared with region growing, watershed, fuzzy-c-means, active contours, this graph-cut approach got a better performance.

Other graph-structures have been used. Based on hierarchical structures, the max-tree graph has been considered convenient for PET image representation. This way, an image is level-set partitioned producing connected-voxel regions for nodes representation which are increasingly ordered and connected, setting up high intensity regions towards the extremes (leaves) of the graph, such as trees. *Grossiord et al.* [60] used the max-tree for lymphoma segmentation with a 92% efficiency for detection, 0.73 sensitivity and 0.99 specificity.

### 1.2.5 Joint approaches

In clinic, PET/CT fusion is widely used to delineate lesions. PET provides sensitive functional tracks while CT gives more accurate location and a better tumor morphology. Some authors tried to reproduce this *modus operandi* adapted for PET/CT multimodal segmentation. In relation with how these co-sources are exploited, methodologies can be classified according to the influence they

provide in the segmentation result. They can guide the whole process for fully co-segmentation or partially at the end of the process by fusing independent segmentations.

*Bagci et al.* [1] present an approach where the random walker algorithm is applied on PET combined with complementary sources such as CT and MRI to model a hybrid segmentation. Image inputs are taken as graphs intra-connected by pixel/voxel neighborhood and inter-graphs connected by spatial correspondence between image matrices to form a hypergraph. Foreground and background seeds are obtained from PET by a threshold protocol on $SUV_{max}$ and projected by a masking to $\Omega$ sources. Random walkers are then individually applied on each data source, whose belonging probability maps are fused to obtain a hybrid segmentation. On different cancer experimentations, results of this method on different multimodal input variations, compared with 40% and 50% threshold methods, region growing, FLAB, show a higher performance. Between the possible combinations that are low to high accuracy ordered (PET only, CT only, MRI only, PET-CT, PET-MRI and PET-CT-MRI), there is a tendency to improve accuracy (0.832, 0.878, 0.881, 0.914, 0.931, 0.934 DSCs) according to quality and quantity of sources injected as input.

*Song et al.* [2] proposed a bimodal methodology based on the graph-cuts for PET and CT segmentation as a minimization problem of a Markov random field model. Multimodal graph structure is similar to that presented by *Bagci et al.* with intra- and inter-edges for hypergraph building. However, *Song et al.* simply presents a two input hypergraph (bi-graph), extended with an extra edge type, called sinking link. This link is used to measure the region term that assign how the intensity of a pixel/voxel fits into an intensity distribution learned from the seed set. Seeds are manually defined in a circular boundary where tumor is inscribed. Whole energy function is composed of a PET and CT respective individual energies equally formed by a region term associated to a boundary term (local intensity gradient from intra edges); and context energy which penalizes PET and CT agreements at labeling. Total energy is iteratively evolving for voxel/node class labeling to find max-flow/min-cut solution. Experimentation with 23 PET/CT with lung tumors and 2 PET/CT head and neck cancer shows that using PET/CT co-segmentation method has an improvement of 41.7% over CT-only and 18.7% over PET-only.

*Ju et al.* [61] suggested to increase the robustness of *Song et al.* method, firstly by automatically initializing seeds on PET. Foreground seeds are selected by classical random walker method whereas background seeds are chosen by a downhill distance surrounding the tumor. Then, energy definition is specified for PET and CT energies, just sharing data and boundary terms but differing at extra terms: down-hill and derivative terms for PET, and shape term for CT (CT data compared with mask of foreground seeds). Results expose a 0.84 DSC versus a 0.76 DSC of *Song et al.* and a 0.74 DSC of *Bagci et al.* on a 18 PET-CT data set of non-small cell lung cancer patients.

### 1.2.6 Discussion

PET accurate segmentation is challenging due to intrinsic image properties such as low spatial resolution and the object of interest quality because tumors tend

to have irregular metabolism and composition. Despite of this heterogeneous composition, PET functional representation provides a good contrast to discriminate tumors which have relatively higher intensities than their background. This main feature allows basic segmentation methods, such as threshold approaches to provide good enough results by just defining a threshold value on the maximal intensity. In spite of noise sensitivity, gradient-based methods try to avoid noise amplification problem by applying denoising pretreatments and then, iteratively handling heterogeneity until converging toward the optimal boundary. Clustering methods show their advantages to fully automate image classification. For instance, fuzzy-c-means allows to analyze fuzzy objects as the tumor-to-background transition zones. Moreover, for more complex algorithms such as graph-based methods, fuzzy-c-means has been adopted as a popular starting method to provide seeds. Graph-based methods demonstrate a versatile framework of intensity interpretation for image graph modeling, and then to be used for segmentation issues. Random walker and graph-cuts show varieties for data processing by respectively direct computation or iterative energy minimization. For its part, max-tree seems to increase PET modeling advantages for tumor regions representation and segmentation facilities. A tumor, represented by high intensities, has a convenient distribution on terminal branches of this tree structure.

In relation with input data quality, multimodality displays better results on co-segmentation approaches. In *Bagci et al.* method, it is shown that accuracy improves with quality and quantity of data input. PET is used for all experimentation tests as seed extractor because of its high tumor sensitivity. However, PET monomodal random walker segmentation has the lowest accuracy between the combinations presented due to heterogeneity and low resolution. MRI only version shows a higher accuracy than CT only. This is directly related with physical bases given that MRI provides a higher contrast between soft tissues. Finally, the PET-CT-MRI input combination furnished the most accurate result, highlighting the relevance of multimodality. *Song et al.* and *Ju et al.* graph-cut based methods also argue that multimodality provides a better performance than individual versions. However, contrary to *Bagci et al.*, these both approaches considered PET monomodal data higher discriminant than CT data. For *Song et al.* approach, CT only version provided a higher accuracy than PET only version, just for regions smaller than 50 $mm^3$ while for larger lesions, PET only version was more accurate as for *Ju et al.* approach on general size experimentation. This contrasted behavior can be determined by ground-truth delineation and tumor types. *Bagci et al.* applied a general disease experimentation while *Ju et al.* and *Song et al.* are mainly specified on lung cancer. *Ju et al.* launched a comparison with an homogeneous data between *Bagci et al.* and *Song et al.* PET/CT approaches, and their approach arguing a higher performance for its method and generally for graph-cut approaches over jointed random walker segmentation. It should be noticed that multimodality for these three approaches remains relevant to improve volume segmentation accuracy.

We can assume, without loss of generality, based on different configurations presented above, that a graph structure provides versatility to be modeled as well as to characterize and weight intensity relationships. Moreover, multimodality increases data richness to be exploited in the graph structure and characteriza-

tion. These both concepts, graph-model and multimodality are selected as the basis to proposed methodologies that allow to improve segmentation accuracy.

# Chapitre 2 : Proposition méthodologique pour la segmentation

La première section de ce chapitre est consacrée à une présentation détaillée des concepts de base utilisés dans la partie méthodologique. Plus spécifiquement, on s'intéresse aux structures à base de graphes comme les structures hiérarchiques qui permettent la représentation d'une image sous forme d'un arbre. En effet, cette configuration permet de mettre en évidence et d'exploiter des informations spatiales et spectrales permettant de structurer une image en régions dont la hiérarchie est définie par rapport aux niveaux d'intensité.

Dans les sections suivantes, nous nous intéressons aux structures hiérarchiques pour proposer trois modèles basés sur des arbre de coupes et arbre des formes dans le contexte de la segmentation tumorale. Nous y introduisons des attributs dont la définition est orientée régions, en mettant en lumière leurs propriétés spécifiques pour la caractérisation des images. Des exemples de traitement basés sur des arbres (filtrage, segmentation) et exploitant les attributs sont alors introduits. Nous nous intéressons au MSER en particulier comme approche de filtrage, ce dernier assurant la stabilité des attributs. Finalement, le marcheur aléatoire est utilisé comme méthode de segmentation, compte tenu du fait qu'elle soit particulièrement adaptée à notre démarche méthodologique.

## Fondements théoriques

La représentation discrète d'une image $I$ peut être interprétée comme étant une fonction définie sur un support $\Omega$. L'ensemble de ses éléments $x \in \Omega$ sont alors associés à une valeur d'intensité $q \in Q$.

Une structure hiérarchique $\Psi$ appelée arbre, est une structure de données qui modélise les informations spatiales et spectrales d'une image, de manière multiéchelle. D'un point de vue global, un arbre est composé de nœuds reliés par des arêtes selon une hiérarchie généralement spatiale, où le niveau le plus bas est la racine et les plus hauts les feuilles.

Soit $\lambda \in \Lambda$ un niveau hierarchique dans l'arborescence $\Psi$. Chaque $\lambda_l$ est lié à l'organisation des nœuds obtenus par la partition d'image, où $l \in \mathbb{N}\,[1, \mathrm{L}]$ fait référence à la numérotation $\lambda$.

Un nœud $N \in \Theta$ est formé par un ensemble de voxels[1] qui sont connectés en

[1] Le terme voxels est ici utilisé car les modèles sont adaptés pour travailler les images 3D. Ce formalisme reste cependant valable en 2D.

considérant une adjacence de voisinage (composantes connexes) dans la partition de l'image (généralement par seuillage).

Ces nœuds $\Theta$ peuvent être liés par leur ordre d'imbrication spatial, des feuilles à la racine $N_{feuilles} \subseteq N_{racine}$, en obtenant ainsi une hiérarchie spatiale enfant / parent selon un ordre de taille croissant.

Deux types de structures hiérarchiques sont ici considérés : l'arbre de coupes et l'arbre de formes.

## Structures hiérarchiques : arbre de coupes et arbre de formes

L'**arbre de coupes** est défini comme étant un ensemble de régions seuillées de l'image $I$. Les régions ou nœuds sont organisées selon leur inclusion spatiale dont la convention de seuil (partition) a la particularité d'avoir un ordre défini par rapport aux intensités $q$, soit $\geq$ pour un max-tree, soit $\leq$ pour un min-tree. Au total, il existe un nombre de partitions égal au nombre de valeurs d'intensité.

L'**arbre de formes** est défini comme étant une variante auto-duale d'un arbre de coupes de part l'utilisation des deux conventions de seuil ($\geq$ et $\leq$) pour le partitionnement de l'image. Cette liaison peut être vue comme une "imbrication isocontour" avec la particularité que les régions sont sans trous par rapport à l'arborescence obtenue, par exemple, avec les arbres de coupes.

Ces deux structures permettent d'organiser les régions de l'image en plaçant celles de haute intensité vers les feuilles de l'arbre. Cette particularité convient pour représenter des objets qui sont considérés comme les maxima locaux de l'image.

## Attributs basés sur les régions et MSER pour assurer la stabilité dans l'evolution d'un attribut

En analysant l'ensemble des voxels inclus dans une région, des **attributs basés sur des régions** peuvent être calculés afin de les caractériser à partir de différentes informations : spectrales, spatiales, géométriques ou hybrides.

Le *maximally stable extremal regions* (**MSER**) est un algorithme permettant d'évaluer les arbres par le biais d'une recherche des nœuds stables par rapport à une métrique. Cette analyse débute au niveau des nœuds extrêmes $N_{feuilles}$ et s'oriente vers la racine $N_{racine}$, via les branches. Un nœud est consideré comme étant stable si, par rapport à un intervalle de nœuds $\pm\Delta$, sa métrique de stabilité $s(h)$ correspond à un minimum local par rapport à un attribut $h$.

## Marcheur aléatoire pour le traitement d'images

La méthode du marcheur aléatoire à été introduite comme méthode de classification adaptable dans le contexte de la segmentation. Elle détermine la probabilité $p$ qu'un marcheur, partant d'un voxel non marqué, atteigne un des points marqués pour définir son appartenance à la classe. Une image $I$ peut être vue comme un graphe connecté, pondéré et non dirigé $G(V, E)$, où chaque voxel $x \in \Omega$ représente un nœud $v \in V$ avec une valeur d'intensité $q \in Q$. Les nœuds $V$ sont liés spatialement par les arêtes $e \in E \subseteq V \times V$. Cette relation non-dirigée entre deux nœuds voisins $v_i$ et $v_j$ est donc exprimée par $e_{ij} = e_{ji}$, et pondérée par une

fonction de poids $w_{ij}$. À partir d'une formulation combinatoire de l'intégrale de Dirichlet et d'une solution harmonique, il a été proposé de rechercher des probabilités sur des nœuds non marqués $p_U$ à partir de nœuds marqués, dont la solution est unique et reproductible. Les principaux avantages de cet algorithme sont sa capacité à obtenir un résultat de calcul direct (non itératif) et à fournir des résultats robustes dans des images dans lesquelles les contours sont partiellement définis (comme les tumeurs par exemple).

# Méthodes développées

Notre travail de recherche se décompose en trois approches méthodologiques qui visent à segmenter des objets caractérisés comme étant des maxima locaux avec des contours flous dans l'image. De plus, ces méthodes sont soumises respectivement à des contextes monotemporel et multitemporel. Elles reposent sur la modélisation d'une image sous forme d'arborescence définie par rapport à des régions spectrales. Cependant, elles diffèrent par la façon dont cette arborescence est exploitée.

## Marcheur aléatoire guidée par des attributs basés sur les régions

Dans cette méthodologie, notre contribution repose sur l'exploitation d'un arbre comme base de traitement et l'extraction des informations enrichies (sous forme vectorielle). Elle permet de procéder à une segmentation par marcheur aléatoire.

Une image vectorielle est définie ici comme une image multicanaux $I_A$ avec voxels $x \in \Omega$ dans un espace matriciel valué par un vecteur $A$ avec des canaux $h$. Le mappage vers un canal $h$ d'un voxel $x$ dans l'espace matriciel de $I_A$, est noté $I_A(x[h])$. Un arbre $\Psi$ est utilisé pour extraire des informations afin de construire $I_A$. Une image d'entrée $I$ peut ainsi être transformée, en passant par un arbre $\Psi$, en une image vectorielle $I_A$ par un protocole d'extraction de l'information souhaité $h_*$ à partir de l'arbre.

Cette étape peut se décomposer en : la création d'un arbre dont le calcul des attributs choisis, le filtrage MSER puis la retro-projection d'attributs en vue de l'obtention d'une image vectorielle. Le MSER est ici appliqué pour ignorer les éléments bruités tout en préservant les nœuds stables en fonction de l'évolution d'un attribut $h$. Un delta $\Delta_h$ est défini comme un intervalle d'évaluation permettant d'appliquer le MSER à un attribut spécifique $h$ ($h$-MSER).

Chaque attribut $h$ présente sa propre évolution tout au long des nœuds qui constituent une branche. Cette évolution n'est pas nécessairement corrélée avec celle des autres attributs. Ainsi, l'arbre initial $\Psi$ est soumis à des évaluations indépendantes et parallèles $h$-MSER afin d'obtenir des arbres filtrés de manière asymétrique $\hat{\Psi}_h$. Chaque arbre filtré $\hat{\Psi}_h$ est soumis à une projection respective de l'arbre $C_N(x[n]) = h_*$ pour remplir un canal $n$ de $I_A(x[n])$, où $h_*$ est l'attribut sélectionné extrait des nœuds stables $\hat{\Theta}$ de $\hat{\Psi}_{h_n}$.

Cette image vectorielle $I_A$ est traitée par l'algorithme classique du marcheur aléatoire dans le contexte de la segmentation d'images. De cette façon, le marcheur est alimenté par des données asymétriques basées sur des régions stables.

Selon une classe de région d'intérêt $c \in C$, des graines voxels $x_{graines}$ sont sélectionnées. Cela peut être fait de manière interactive en sélectionnant manuelle-

ment les voxels qui sont des membres fortement probables d'une classe $c$ ; ou en appliquant des algorithmes automatiques. Dans le cas de classes binaires, où une classe d'objet $c$ présente un contraste relativement élevé avec son arrière-plan, les approches par seuil peuvent être utilisées pour simplifier la sélection des graines.

Enfin, à partir du nœud graines $v_{graines}$ représentant les classes $C$, le marcheur aléatoire vectoriel est appliqué lors de la phase de segmentation.

## Filtrage par le couplage des arbres pour la segmentation

Dans la deuxième approche que nous avons proposée, nous nous sommes intéressés à la problématique de la segmentation des images multitemporelles, en essayant de résoudre le problème de recherche des similitudes entre les images. Cette absence de correspondance spatiale est abordée par le biais de la modélisation d'une image sous forme d'arbre permettant de remplacer le paradigme voxel à voxel pour celui du nœud à nœud.

A partir d'une image $I_{\mathrm{A}}$ et d'une image $I_{\mathrm{B}}$, des structures hiérarchiques $\Psi_{\mathrm{A}}$ et $\Psi_{\mathrm{B}}$ sont construites. La nature de $\Psi$ dépend de sa pertinence pour la modélisation des objets d'intérêt où la convention de partition de $\Psi$ (arbre de coupes, arbre de formes,...) doit garantir que l'objet d'intérêt est défini dans l'arborescence comme un nœud $N$.

Les attributs pertinents $h$ sont calculés sur les nœuds $\Theta$ afin de les caractériser. Ces valeurs généralement scalaires sont normalisées puis stockées dans un vecteur d'attribut $A$. Leur sélection dépend de leur aptitude et pertinence à décrire l'objet d'intérêt dans les deux ensembles de données.

Les régions $\Theta_a$ et $\Theta_b$ des deux arbres $\Psi_{\mathrm{A}}$ et $\Psi_{\mathrm{B}}$ sont mis en correspondance afin de déterminer les plus similaires. Leur niveau de similarité est mesuré par une distance entre les attributs des nœuds, avec une valeur de tolérance $\varepsilon$. Les régions dont la distance est inférieure à $\varepsilon$ sont alors considérées comme similaires et stockées en tant que candidats potentiels pour l'appariement. Par la suite, une phase de sélection de candidats est appliquée en considérant deux aspects : déterminer le meilleur candidat disponible et gérer le problème de croisement hiérarchique.

Entre les candidats possibles d'un noueds $N_i$, celui qui a la distance $D$ la plus basse est défini comme le meilleur candidat disponible.

Un problème de croisement hiérarchique se produit lorsque la hiérarchie, au couplage potentiel, n'est pas respectée. Si le meilleur candidat respecte le non-croisement de hiérarchie, il est considéré couplé à son nœud et enlevé des candidats au couplage pour les autres nœuds. Dans le cas contraire, ce candidat est effacé de la liste de candidats du nœud lui-même et leur meilleur candidat suivant est évalué jusqu'à épuisement. Cette sélection de candidats est complétée de manière itérative pour tout l'ensemble des nœuds, soit les nœuds $\Theta_A$ ou $\Theta_B$, afin d'obtenir un couplage $\xi(N_a, N_b)$ muni de son poids $sum = \sum exp(-D(N_a, N_b))$. L'ordre de sélection des candidats sera évalué hiérarchiquement en laissant tous les nœuds $N_{\mathrm{A}}$ et $N_{\mathrm{B}}$ choisir leurs candidats en priorité. Le couplage $\xi$ présentant le plus haut $sum$ est considéré comme étant le couplage final $\Xi$.

Un protocole de segmentation est ensuite appliqué afin d'identifier les nœuds appartenant à l'objet d'intérêt. Ces nœuds sont reprojetés dans leur espace de support d'image initial pour récupérer le résultat de la segmentation.

## Filtrage par vote majoritaire MSER et marcheur aléatoire sur un graphe hiérarchique

La troisième méthode que nous avons proposée est une extension de la première. L'arborescence est ici directement impliquée en tant que structure de données pour l'application d'un algorithme de segmentation.

Un arbre $\Psi$ est construit à partir d'une image $I$. L'algorithme de stabilité, par vote majoritaire MSER, est appliqué pour détecter et pour élaguer les nœuds instables vis à vis de l'évolution des attributs $h$ à travers les branches de l'arbre.

Les graines $v_{graines}$ dans l'espace image sont projetées dans l'arborescence en tant que étiquettes pour les nœuds graines $N_{graines}$. Les étiquettes associées à une classe sont propagées par voie hiérarchique dans tous les descendants d'un nœud.

Deux classes sont ici définies : une classe active et une classe passive.

Pour une classe active $c_*^+$, tous les descendants d'un nœud $N_i$, c'est-à-dire un sous-arbre $\psi(N_i)$ avec la racine $N_i$, sont des membres impératifs de la classe $c_*^+$.

Une classe de nœuds passifs $c_*^-$, généralement utilisée pour représenter l'arrière-plan, fournit à leurs nœuds membres la particularité d'une hiérarchie négative dans laquelle, à partir d'un nœud $N_i$, tous les nœuds ascendants dans sa branche, bien qu'il n'y ait pas des nœud d'une classe active, sont étiquetés comme membres impératifs de la classe $c_*^-$.

A partir de cette propagation, l'arbre est orienté vers la racine par le fond $(c_{bg}^-)$, vers les feuilles par les objects $(c_*^+)$, en définissant des zones d'incertitude intermédiaires $\kappa$. Ces sections de nœuds incertains $\kappa$, avec des attributs basés sur les régions en tant que descripteurs, sont impliquées dans une méthode de marcheur aléatoire vectoriel pour compléter l'étiquetage de l'arbre et pour réaliser la segmentation.

Ainsi, en considérant un ordre hiérarchique où la racine $N_{racine}$ est le niveau le plus élevé et les feuilles $N_{feuilles}$ sont le niveau le plus bas, l'extremum $N_i(c_*^+)$ de l'arbre $\Psi$ correspond à l'objet segmenté pour chaque classe positive $c_*^+$.

Des informations contextuelles provenant d'autres images peuvent être injectées dans l'arbre en profitant de leur correspondance spatiale. De cette manière, l'information considérée comme étant pertinente est exploitée pour extraire des attributs $h$ qui vont intervenir lors du filtrage par vote majoritaire MSER et l'étape de segmentation basée sur l'exécution d'un marcheur aléatoire sur l'arbre.

# Chapter 2

# Methodological proposals for segmentation

## Contents



## 2.1 Theoretical basis

In order to solve the PET segmentation problem, graph-based structures and multimodality have been selected as the core of our approaches. This section is dedicated to a detailed explanation of these basic concepts and justify their advantages. Specifically about graph-based structures, hierarchical structures are selected as graph models for image representation. Indeed, this configuration allows to extract spatial and spectral image information, obtaining image regions with a hierarchical order on intensity levels. Two classical hierarchical models are presented in Section 2.1.1: the component-tree and the tree of shapes. In Section 2.1.2, region-based attributes are introduced, and we shed light on their specific properties for image characterization. In Section 2.1.3, some examples of hierarchical graph processing (filtering, segmentation) by region-based attributes are introduced to reinforce their understanding. In Section 2.1.4, a particular filtering, related with attribute stability, is presented. Finally, in Section 2.1.5, random walker is explained in detail because of its specific use in proposed approaches.

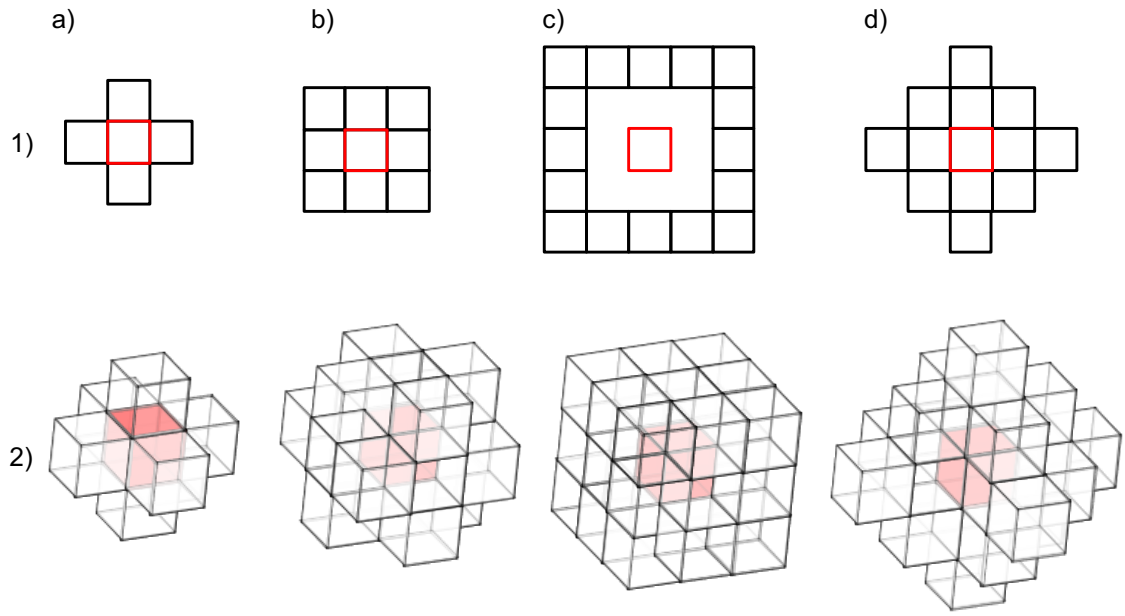


Figure 2.1: Graphical examples of neighboring adjacencies $nb$ in 1) 2D: 4, 8, 16 with $\Delta = 1$, 12 star; and 2) 3D: 6, 18, 26, 24 star; where red square/cube is the pixel/voxel whose $nb$ is represented.

### 2.1.1 Hierarchical structures: Component-tree and tree of shapes

An image $I$ can be physically described as a spatial representation of a physical quantity. A discrete representation of an image $I$ can be defined as the mapping function of the support $\Omega$, the set of its elements (2D pixels/3D voxels) $x \in \Omega$, related with an intensity value $q \in Q$, $I : \Omega \rightarrow Q$.

From another point of view, based on mathematical morphology, an image $I$ can be reconfigured by considering its voxels $\Omega$ with respect to its spectrality (intensity levels $Q$) and/or its spatiality (distribution at matrix lattice).

A hierarchical structure (also known as tree), denoted by $\Psi$, is a data-structure that models the spatial and spectral information of an image in a multiscale way. The levels of the tree define a hierarchy. A tree is composed of nodes, linked by edges according to a commonly spatial hierarchy where the lowest level is the root and the highest the leaves.

Let $\lambda$ be a hierarchical level of the tree $\Psi$. Each $\lambda_l$ is related to the organization of nodes obtained by image partition, where $l \in \mathbb{N}[1,\mathrm{L}]$ refers to $\lambda$ numbering and L is the highest level.

A partition refers to the splitting of pixels/voxels on foreground and background binary labels, denoted by $x_{fg} \in \Omega_{fg}$ and $x_{bg} \in \Omega_{bg}$ where $\Omega_{fg}, \Omega_{bg} \subseteq \Omega$, provoked by a spectral condition at intensity $q$. In an image, the number of partitions $n$ is related with its spectral resolution where $n$ can be less or equal than $|Q|$ according to partitioning protocol and $|Q|$ denotes cardinality of $Q$.

Voxels $\Omega_{fg}$ of each partition are then evaluated to form nodes $N$. A node $N \in \Theta$ is formed by a set of pixels/voxels in the subset $\Omega_{fg}$ that are connected considering a neighborhood adjacency. Let $x_i, x_j, x_k$ denote three pixels/voxels of $\Omega_{fg}$. $x_i, x_j$ are adjacent if they are incident to a neighborhood relationship according to the image lattice (Figure 2.1). Let $x_i$, $x_j$ and $x_j$,$x_k$ be pairs of adjacent nodes

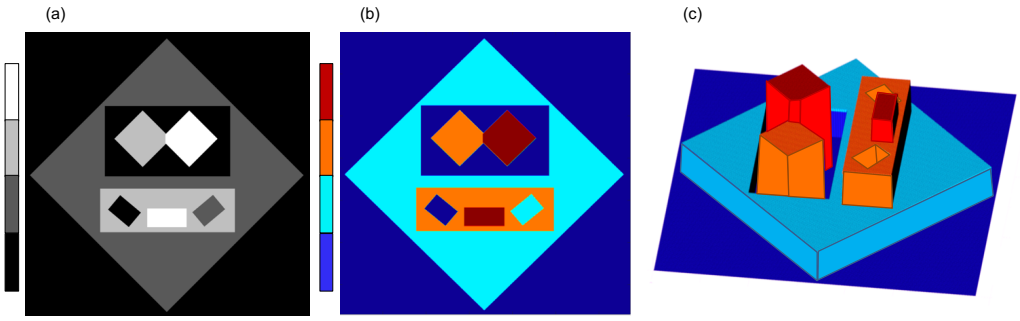


Figure 2.2: (a) 2D image with 4 gray levels, (b) colormap of (a) and (c) 3D relief of (b).

while $x_i$ and $x_k$ are non-adjacent ones. By transitive property, $x_i$, $x_j$ and $x_k$ are connected between them. A node is known as a maximal connected component with respect to an adjacency. As origin signature, denoted by $\vartheta$, nodes can be related with the intensity level $q$ of their partitioning condition by $N(\vartheta) = q$.

These nodes $\Theta$ can be linked by their spatial nesting order, from the leaves to the root $N_{leaves} \subseteq N_{root}$, thus obtaining a spatial hierarchy (child/parent) on an increasing size order, finally completing tree building.

The operators are used $N_p(N_i)$ to call the parent and $N_c(N_i)$ for the children, and their respective number of elements $n_p(N_i)$ and $n_c(N_i)$ of a node $N_i$. The number of children $n_c(N_i) \in \mathbb{N}$ can be larger or equal to 1 where just the leaves have the particularity of $n_c(N_{leaves}) = 0$. Whichever node can just have one parent $n_p(N_i) = 1$, excepting the root $n_p(N_{root}) = 0$. A node ancestor $N_{anc}$ of $N_i$, denoted as $N_{anc}(N_i)$, is defined as a node belonging to the nested node sequence $\Theta_{anc}(N_i)$ from $N_p(N_i) \subseteq N_{root}$.

A branch section of a tree, denoted by $\varphi(N_i, N_{anc}(N_i))$, is defined as a nested node sequence from a beginning $N_i$ to one of its ancestors $N_{anc}(N_i)$. A branch section example can be expressed by a node $N_i$ and its parent $\varphi(N_i, N_p(N_i))$, whereas for a complete branch section by $\varphi(N_{leaves}, N_{root})$.

Hierarchical graphs have the general structure previously explained. The main difference between their varieties refers to the partition protocol used for the conception of image regions $\Theta$. In this work, the component-tree and the tree of shapes are considered as potential structures to be exploited. These both hierarchical structures assume that an image $I$ can be seen such as a topological relief (Figure 2.2), where intensity values are height levels representing elevation and depth respectively. However, the main difference between them is established by how their levels are determined and interpreted.

### Component-trees

The component-tree $\tau$ of $I$, introduced by *Salembier et al.* in 1998 [62], is defined as a directed threshold collection of $I$, whose binary regions (connected components) are hierarchically organized by their spatial inclusion. At each $q$ a threshold convention ($\geq$ for a max-tree or $\leq$ for a min-tree) is applied to produce a binary partition at $\lambda_q$ level. If pixels/voxels comply with the threshold agreement, they are marked to be foreground, otherwise as background. In total, there is a number of partitions equal to the number of intensity values in

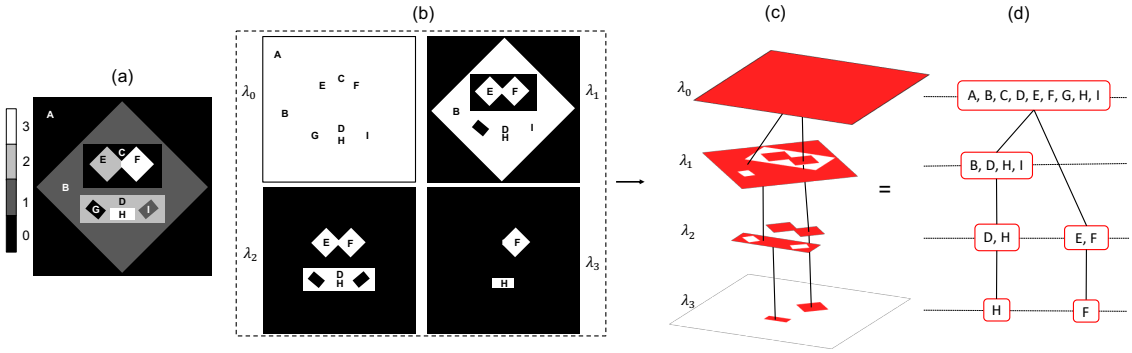


Figure 2.3: Graphical process of max-tree building. 2D image with $[0-3]$ gray levels (a) with its level-set partition ($\geq$) of (b). (c) Graphical representation of inclusion linking of a max-tree (d) at levels $\lambda_q$.

the interval $[q_{min}, q_{max}]$, where $q_{min}$ and $q_{max}$ denote minimum and maximum intensities of an image. Following the spatial inclusion order on $\lambda_l$, $l$ is directly related to $q$ due to sequential thresholding, thus $L = |Q|$. In Figure 2.3 an example of max-tree building is presented. From $N_{leaves}$ at $\lambda_3$, nodes have a increasing size (greater regions) until the $N_{root}$ which is composed of all pixels. The image regions, denoted by C and G for example, are not node represented under this partition type because of having a negative contrast region-to-background (a spectral hole), $d$(C,B) $= 0-1 = -1$ and $d$(G,B) $= 0-2 = -2$ where $d(a,b)$ is the subtraction between region $a$ and $b$.

Several algorithms have been proposed to improve component-tree computing presented by *Salembier et al.*, from the root to the leaves having a complexity of $O(n \times L + m)$ where $m$ denotes the number of operations, $n$ denotes the number of elements and $L$ is the intensity level interval $q_{max} - q_{min}$. *Najman et al.* [63] proposed an algorithm, from the leaves to the root, to compute the component-tree in quasi-linear time $log(n)$.

### Tree of shapes

The tree of shapes $\phi$ of $I$, introduced by *Monasse et al.* [64] in 2000, is defined as a self-dual version of the component-tree using both threshold conventions ($\geq$ and $\leq$) for image partitioning. Indeed, due to regional composition, spatial inclusion does not necessarily have a spectral order for $\lambda_l$ sequence because of being composed of nodes from both min/max conventions (Figure 2.4). This linking can be seen as an "isocontour nesting" with the particularity of non-holed regions compared with component-tree. Indeed, holes (C, G, I nodes for example) are added as regions which are provided by threshold convention $\leq$.

In the "Fast Level Line Transform" algorithm [65], a max-tree and a min-tree are computed, merging both trees to built the tree of shapes, with the main constraint of two tree time computing. In another attempt, the "Fast Level Set Transform" [66] is used to detect local maximum levels (leaves) in order to apply a region-growing to build the branches to merge until the root. In 2013, *Géraud et al.* [67] present an algorithm to compute the tree of shapes in quasi linear time, based on an algorithmic "canvas" [68], for computing nodes of dual partition ($\geq$ and $\leq$), recording their inter-node linking, and then, render a tree as these data

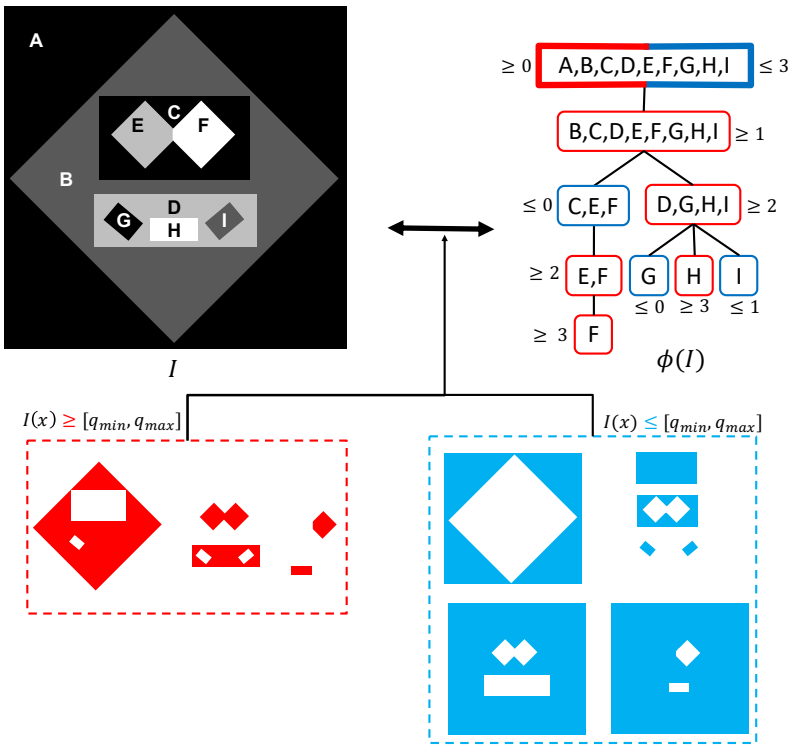


Figure 2.4: Isocontour partitioning and linking of $I$, for tree of shapes $\phi(I)$ building. Nodes coming from $\geq$ (blue) and $\leq$ partitions of $I$, are inscribed in respective dotted boxes and graphically represented in the tree of shapes $\phi$.

interpolation. This is the retained algorithm that is used for this work experimentation, either for max-tree than for tree of shapes computing.

### 2.1.2 Region-based attributes concept

The nodes of a tree can be valued by specific features to describe a region itself. The concept of region-based attribute [69] was introduced as a feature used to characterize spectral image regions. In essence, a region on an image can be from just one pixel/voxel, until the whole image. The term "region" gives directly a notion of connectivity on a spatial framework. By analyzing the set of pixels/voxels included in a region, region-based attributes can be spectral, spatial, geometrical or hybrid (Table 2.1).

Spectral attributes are built from intensity levels $q$ of pixel/voxel $x$ related to the node $N$. Minimal intensity $N(x(q_{min}))$, maximal intensity $N(x(q_{max}))$, inten-

| Type | Region-based attributes |
|---|---|
| Spectral attributes | $x_{q_{min}}$, $x_{q_{max}}$, $x_{q_{mean}}$, $x_{q_{median}}$, $\text{cont}_{\text{int}}$, $\text{cont}_{\text{ext}}$ |
| Spatial attributes | Area, $\text{area}_{\text{cont}}$ barycenter coordinates |
| Geometrical attributes | BA, compactness, complexity |
| Hybrid attributes | spectral volume, volumetric contrast |

Table 2.1: Examples of region-based attributes.

sity mean $N((\Sigma_a^{n(x)} x_a(q))/n(x))$ and intensity median $N(x(q_{median}))$ are some basic examples. Contrast attribute gives a spectral difference notion. Two types of contrast are introduced: internal contrast (spectral height) $\text{cont}_{\text{int}}$ and external contrast $\text{cont}_{\text{ext}}$. Internal contrast refers to the subtraction between maximal and minimal intensities from node voxels by $\text{cont}_{\text{int}} = x(q_{max}) - x(q_{min})$. External contrast is the intensity subtraction between two neighbor nodes $N_i$ and $N_j$ (for a child-parent relation) by $\text{cont}_{\text{ext}} = N_i(\vartheta) - N_p(N_i)(\vartheta)$, where, making a recall, $\vartheta$ is the characteristic intensity $q$ for a node $N$.

Spatial attributes focus on space of the object itself such as area (size), $\text{area}_{\text{cont}}$ as $|N_i(\text{area}) - N_p(N_i)(\text{area})|$ and barycenter coordinates to provide a relative distribution inside a global framework.

Geometrical attributes rely on geometry definitions. Boundary area (BA) is the number of pixels/voxels of the boundary of a node (perimeter/surface) according to an adjacency. Circularity/sphericity (compactness or roundness) is the comparison between a node and a circle/sphere by a BA/area ratio. Compared with compactness, complexity (elongation) is a more sensitive parameter with respect to boundary area analysis by an inverse ratio $\text{area}/\text{BA}^2$.

Finally, hybrid attributes are computed by merging two attribute types such as spectral volume (area weighted by intensity values) and volumetric contrast (volume weighted by $\text{cont}_{\text{ext}}$), which both are a spatial-spectral combination.

### 2.1.3 Hierarchical graphs for image processing

Modeling an image as a hierarchical graph can be convenient to different ends. Pruning a tree removing nodes seen from the tree, can be considered as a filtering process and from the pruned nodes as a segmentation process. Taking into account node characteristics, a node classification process can be applied.

**Filtering based algorithms** For example in the case of filtering, a tree can be used to remove some types of noise. For "salt and pepper" noise, low and high intensity small stains are present on the image. Thus, a tree of shapes is a good choice to represent both spectral noise modes. Area filtering experimentation is applied on a 2D image on gray levels (Figure 2.5), pruning nodes whose area is lower than a given $\varepsilon_{area} \in \mathbb{N}$.

In Figure 2.5, noisy strains are removed as well as relevant data complying with the threshold. In 1.b), eyes and eyebrows as some details in the camera are cut off while in 1.c) the cameraman face and several objects are disfigured due to overmuch $\varepsilon$ selection. On the other hand, contrast filtering can be here appropriate since noise has high contrast. Contrast filtering is applied on $I$ decreasingly varying $\varepsilon_{contrast}$. In 2.a), as $\varepsilon_{contrast}$ is not discriminative enough, on the sky which has lower contrast with high intensities, white strains remain while in darker zones such as the grass, some black strains are present. In the case of using a max-tree (min-tree) image representation for this noise filtering, just "salt (pepper)" noise is removed.

A second filtering step by a min-tree (max-tree) can be applied, to compensate the remaining inverse noise. However, there is a higher computing cost by two tree filtering processes. Moreover, it should be noted that results from max-min-tree bi-filtering compared with tree of shapes filtering using same area/contrast


Figure 2.5: Filtering experimentation by a tree of shapes $\phi$ and a component-trees (max-tree $\tau^{max}$ and min-tree $\tau^{min}$). 1.a) 2D image $I$ with a $257\times258$ pixels spatial resolution, $[0,255]$ gray levels and "salt and pepper" noise of 0.02 density. 1.b) and 1.c) filtered images reconstructed from $\phi(\Theta(\text{area})) > \varepsilon_{area}\{20,600\}$. 2.a) and 2.b) filtered images from $\phi(\Theta(\text{cont}_{\text{ext}})) < \varepsilon_{contrast}\{100,140\}$. 2.c) Fused area & contrast bi-conditioned filtered image from $\phi(\Theta(\text{area})) > \varepsilon_{area}\{10\}$ & $\phi(\Theta(\text{cont}_{\text{ext}})) < \varepsilon_{contrast}\{140\}$. 3.a) and 3.b) filtered images from $\tau^{max}(\Theta(\text{area})) > \varepsilon_{area}\{20\}$ and $\tau^{min}(\Theta(\text{area})) > \varepsilon_{area}\{20\}$, respectively, and 3.c) image result of these last two filtering processes sequentially applied in their appeal order.

parameters, are not totally similar because of their different inclusion order.

**Segmentation based algorithm** Components trees have already been used as part of the segmentation. In 1999, *R. Jones* [70] proposed to use a component-tree to segment the face of the Mona Lisa and studying the evolution of area attribute through its branches. *Naegel et al.* proposed an interactive method of segmentation of moles based on the use of attribute vectors [71]. They combined

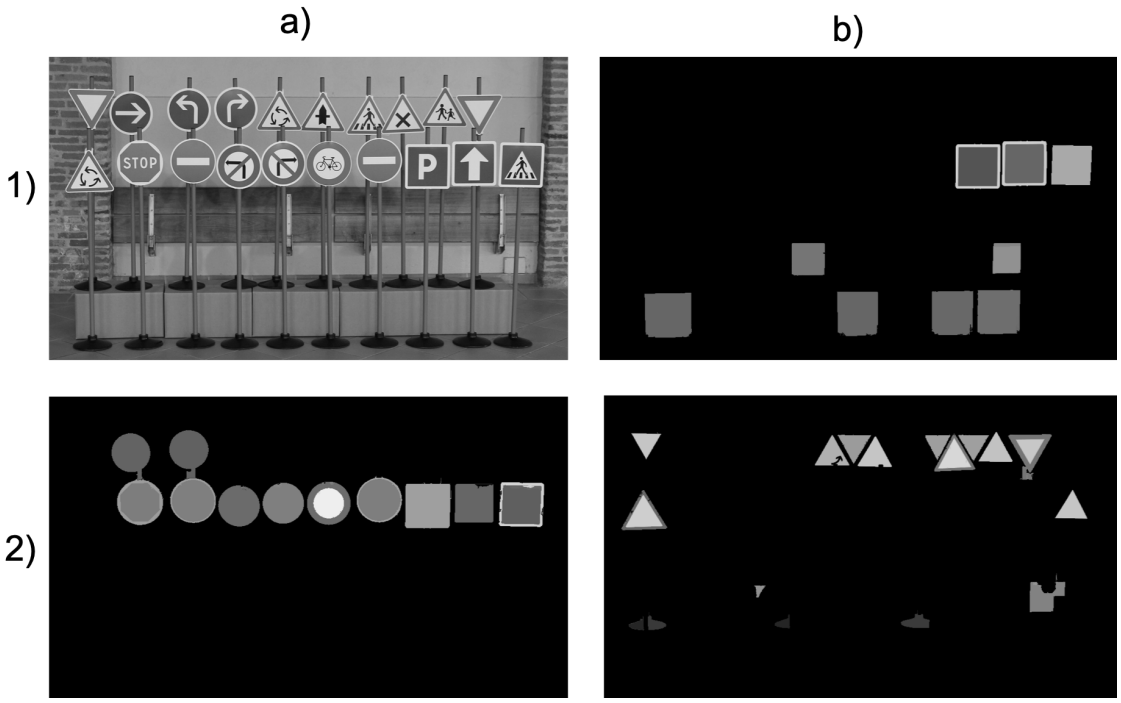


Figure 2.6: Geometric figure segmentation by a tree of shapes $\phi$. 1.a) 8 bits gray level $I$. 1.b) $I'$ with a 0.95 $\varepsilon_{areabox}$ and a 0.9 $\varepsilon_{sqaxes}$ for square segmentation. 2.a) $I'$ with a 0.9 $\varepsilon_{circ}$ for circle segmentation. 2.b) $I'$ with a 0.2 $\varepsilon_{areabox}$ and a 0.1 $\varepsilon_{pertri}$.

the information on the area, contrast and compactness of the elements of interest.

In relation with geometrical figures, such as circles, rectangles and triangles, their representative features can be adapted on region-based attributes to be tracked and segmented from a tree (Figure 2.6).

Circles are characterized by their radius relation with contour and area. From a given area, a circle is the shape which provides the lowest possible perimeter (boundary area) by the notion of compactness. This association can be reproduced by computing the ideal circle perimeter from a node area. The ratio between ideal circle perimeter and node perimeter can give a normal metric interval $[0,1] \in \mathbb{R}$ to evaluate circularity, where 1 means a perfect circularity. In Figure 2.6.2.a) some squares are also segmented on this circle segmentation due to cropped zones reducing area, increasing perimeter to simulate a false circle under this discrimination rule.

For rectangle segmentation, a bounding box, defined as supremum and infimum of the points in the node, can be computed to obtain the maximal rectangle covering the node. This box area can be equal or higher compared with its node area. Analogically, this allows one to measure how a node fills in its bounding box. This ratio is represented on a normal metric interval from 0 to 1, where 1 refers to a perfect rectangularity.

For a square shape (a rectangle type), comparison between both $x$ and $y$ bounding box axes can be added to measure their similarity preservation $|1 - x/y| < \varepsilon_{sqaxes}$ where 0 is a perfect square (Figure 2.6.1.b).

For triangle segmentation, it can be spatially identified by half bounding box area. Based on its geometry, for an equilateral triangle, perimeter is equal to three

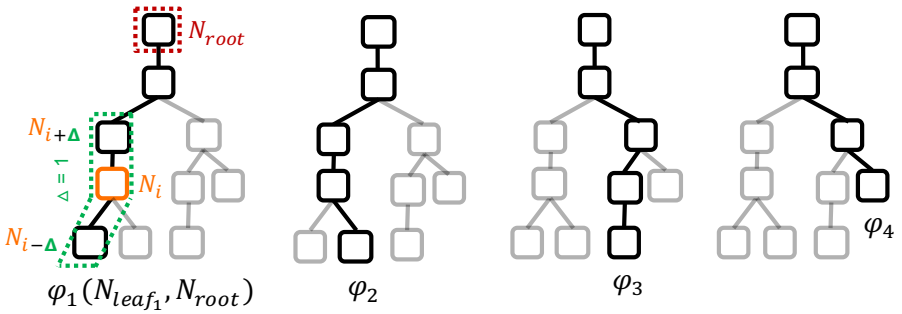


Figure 2.7: MSER paths for each $\varphi_j$ branch. Respective $j$ branch nodes are enhanced with black color whereas non branch nodes are set on gray. The left tree-graph is used to graphically define important terms for MSER processing: current node $N_i$ at branch iteration $i$ in orange, $\pm\Delta$ range in green, and $N_{root}$ in red. The four tree-graphs represent the $\varphi_1, \varphi_2, \varphi_3$ and $\varphi_4$ local branch MSER $j$ iterations to be applied. A node can belong to more than one $j$ branch. This characteristic has a tendency to increase going down level at the tree. Therefore in the same attribute analysis, multiple MSER are applied on a node $N_i$ which are equivalent to branches this node belongs to. If a $N_i$ is stable for at least one $j$ path, it is considered generally stable for the evaluated attribute $h$.

times side length. Thus for a node, its perimeter is compared with three times the base of square bounding box, supposing triangle inclination is 0°, conserving an absolute difference lower than a $\varepsilon_{pertri}$.

In Figure 2.6.2.b), most part of triangular shapes are segmented including those from background. Triangles with partial definition (overlapped, cropped or unclosed) can not be detected because of not complete representation.

### 2.1.4 Maximally Extremal Stable Regions for tree filtering

The Maximally Stable Extremal Region (MSER) is a stability algorithm for tree structures that looks for stable regions in the image. This analysis begins from extremal nodes $N_{leaves}$ (leaves of $\Psi$), up to the root $N_{root}$, through the branches (Figure 2.7).

Area-based stability was presented by *Matas et al.* [72], where a stability state of a node is achieved, if and only if the estimator:

$$s_{\text{area}}(N_i) = |N_{i+\Delta} \setminus N_{i-\Delta}|/|N_i| \tag{2.1}$$

has a local minimum at $i$ where $\Delta \in \mathbb{N}$ denotes node interval. The ratio with $|N_i|$ allows one to normalize area subtraction in a comparable scale given that it is about an incremental attribute.

We generalize this metric to be used for any region-based attribute type as follows:

$$s_h(N_i) = \sum_{j=1}^{2*\Delta} |N_{i+1+\Delta-j}(h) - N_{i+\Delta-j}(h)| \tag{2.2}$$

where this summation represents a sequential subset distance recovery at $\pm\Delta$

range. If $h$ is an incremental attribute as $h$ = area, based on equation (2.1), equation (2.2) is divided by $|N_i(h)|$.

### 2.1.5 Random walker in image processing

Random walker (RW) algorithm was introduced by *Grady* [73] as an image segmentation approach. Automatically defined or user-provided seeds are used to guide object image segmentation. To start the random walker process, unseeded pixels/voxels are labeled by the shortest weighted path towards labeled marks. This solution is equivalent to the one adopted to solve the Dirichlet problem [74]. The analogy of an image graph with an electric circuit allows to model the solution of combinatorial Dirichlet problem by the distribution of their electrical potentials.

The output of the algorithm is expressed by a map of probabilities $P^s$ of belonging to each class $s$. Thus, each pixel/voxel $x_i$ has a vectorial $p_i^s$. Pixels/voxels are then assigned as members of the class $s$ with the highest probability value $p^s$.

In the next subsection the method is explained in two parts: graph weights computing and the solution for equation system.

#### Graph weights computing

An image $I$ can be modeled as a connected undirected weighted graph $G(V,E)$, where each pixel/voxel $x \in \Omega$ represents a node $v \in V$ with an intensity value $q \in Q$. Nodes $V$ are spatially binded by edges $e \in E \subseteq V \times V$. This relationship between two neighbor nodes $v_i$ and $v_j$ is $e_{ij}$ ($e_{ij} = e_{ji}$) and weighted by a Gaussian function:

$$w_{ij} = exp(-\beta|x_i - x_j|^2), \tag{2.3}$$

where $\beta$ is the only free parameter to control weighting interpretation. The square gradient $|x_i - x_j|^2$ is considered useful to be normalized before application of equation (2.3). Color images, or other general vector valued images, can be used here by replacing the type of gradient $\|x_i - x_j\|^2$.

The degree of a node is defined as the weight mass of $w_{ij}$ summation of all incident edges $e_{ij}$ to the node $v_i$, $d_{ij} = \sum w_{ij}$ for all incident edges $e_{ij}$.

#### Solution based on Dirichlet problem

The weighted graph $G(I,E)$ can be expressed here on a positive semidefinite combinatorial Laplacian matrix [73] such as:

$$L_{ij} = \begin{cases} d_i & \text{if } i = j \\ -w_{ij} & \text{if } v_i \text{ and } v_j \text{ are adjacent nodes,} \\ 0 & \text{otherwise,} \end{cases} \tag{2.4}$$

by representing the relation between $v_i$ and $v_j$ with their respective $w_{ij}$ and having in the main diagonal the $d_{ij}$ degree.

From combinatorial formulation of the Dirichlet integral:

$$D[p] = \frac{1}{2} p^T L p, \tag{2.5}$$

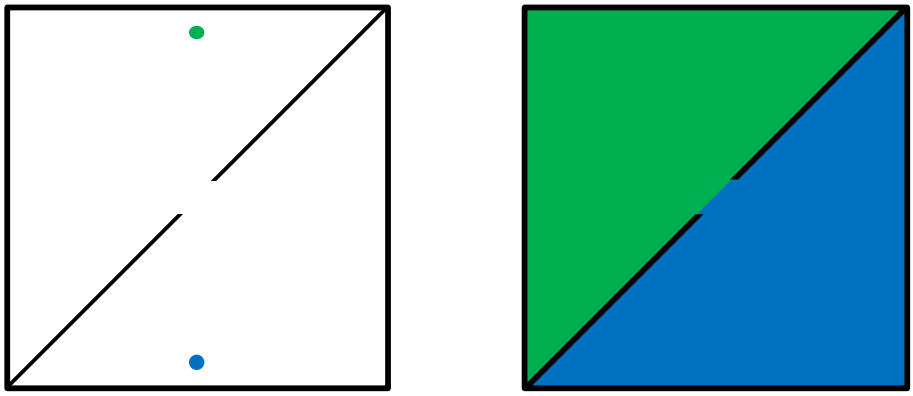

Figure 2.8: Example of two class random walker segmentation on a 2D image (expressed by green and blue colors) with discontinuous inter-border and symmetric size/location seeds.

a harmonic solution has been proposed by *Grady* to find probabilities on non-marked nodes from marked ones, minimizing $p$.

Nodes $V$ can be separated in marked nodes $V_M$ and unmarked nodes $V_U$ where $V_M \cup V_U = V$ and $V_M \cap V_U = \emptyset$. Ordering these subsets, equation (2.5) becomes:

$$D[p_U] = \frac{1}{2}\begin{bmatrix} p_M^T p_U^T \end{bmatrix}\begin{bmatrix} L_M & B \\ B^T & L_U \end{bmatrix}\begin{bmatrix} p_M \\ p_U \end{bmatrix} = \frac{1}{2}\left(p_M^T L_M p_M + 2p_T^U B^T p_M + p_U^T L_U p_U\right), \quad (2.6)$$

where $p_M$ and $p_U$ are the potentials of marked and unmarked nodes.

The differentiation of $p_U$ from $D(x)$ on critical points is applied by:

$$L_U p_U = -B^T p_M, \quad (2.7)$$

where $p$ probabilities can be decomposed in a single label vector $p^s$, where $s$ is a class $s$. By circuit superposition theory, summation of $p_i^s$ for each $v_i$ of all classes $s$, must be 1. Therefore, for markers, according to their respective class, $p_M^s$ is set at 1, other classes at 0.

All unmarked pixels/voxels are guaranteed to be connected to a marked point with the same label. Therefore, it is not possible to have isolated regions. The solution is unique and reproducible. The main advantages of this algorithm is its ability to have a direct computation result (non-iterative) and be performing face to partially defined borders (Figure 2.8).

In our work, this random walker algorithm is used to exploit tree extracted information in the image graph or adapted to be directly applied on the tree structure.

## 2.2 Developed segmentation methodologies

Our research work consists in three methodologies that are based on the modeling of an image as a tree with spectral regions as nodes that are inter-linked by an inclusion relationship. However, they differ in the way this tree structure is exploited.

In the first approach (Section 2.2.1) the tree is used to process and extract information dedicated to feed a random walker segmentation. A set of region-based attributes is used to characterize tree nodes and then project them into the image space for building a vectorial image. A random walker guided by vectorial tree data on image lattice is used to label voxels for segmentation.

In the second approach (Section 2.2.2) we focused on removing spatial correspondence dependency by changing voxel-to-voxel for node-to-node paradigm. Images are then modeled as tree structures, whose nodes are involved in an attribute distance comparison to match those assumed similar whereas discarding the others. A segmentation protocol can be then applied to detect our object of interest between matched regions.

In the third approach (Section 2.2.3) that is an extension of first one, the tree is directly involved as the data-structure for algorithm application. A tree structure is built from an image. A node stability algorithm is applied to detect and prune unstable nodes according to the evolution of region-based attributes through tree branches. Seeds from matrix space are projected into the tree to assign node seed labels and propagate them by hierarchy until defining uncertain branch sections between foreground to background nodes. These uncertain node sections, with region-based attributes as descriptors, are involved in a vectorial random walker method to complete tree labeling and build the segmentation.

## 2.2.1 Random walker in a vectorial image

Hierarchical structures and random walker method have been introduced in sections 2.1.1 and 2.1.5. A hierarchical structure is an image model that allows one to process image data in a spectral–spatial domain that provides image regions linked and ordered by a spatial inclusion relationship. Random walker method is a well-known multi-class image processing segmentation method, where unmarked voxels are related with marked ones by the shortest path paradigm.

In this work, we propose to merge these approaches for exploiting their advantages in order to build and characterize image regions on a tree and project this information to feed random walker algorithm. Image gray levels are the most used metric to weight pixel/voxel difference in the matrix space. However, derived scalar or vectorial information can be computed to describe a pixel/voxel. Thus, a vectorial descriptor can be adapted in weighting function (equation 2.3) to apply a vectorial difference.

Our contribution concerns the image matrix space, building a pixel/voxel descriptor based on region-based attributes $A$ at nodes $\Theta$ to represent a vectorial image $I_A$.

This procedure is divided into two main steps: vectorial image computing by the tree and random walker segmentation in this vectorial image.

### Vectorial image computing

A vectorial image is here defined as a multichannel image $I_A$ with pixels/voxels $x \in \Omega$ in a matrix space that is valued by a vector $A$ with channels $h$. A general mapping to a channel $h$ of a pixel/voxel $x$ into the matrix space of $I_A$, is denoted by $I_A(x[h])$. A tree structure $\Psi$ is used to extract information to build $I_A$. An

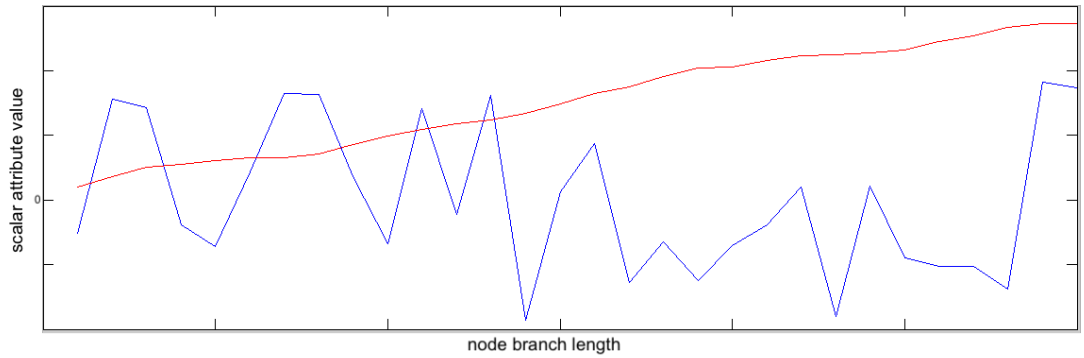


Figure 2.9: Behavior of attributes along a branch. Area (red) as incremental attribute, and $\text{area}_{\text{cont}}$ (blue) as non-incremental attribute, both standard profiles at branch $\varphi(N_{leaf_i}, N_{root})$ of an image $I$, where $i$ refers to an arbitrary $N_{leaf}$.

input image $I$ can be thus transformed, passing through a tree $\Psi$, to a vectorial image $I_A$ by an extraction protocol on region-based attributes.

This transformation can be divided into three steps: tree building, MSER filtering and attribute projection.

A tree $\Psi$ is built on an image $I$ and region-based attributes are computed and stored on $A$ vector on its nodes $\Theta(A[h])$.

### MSER filtering

MSER method is here applied to discard noisy elements whereas preserving stable nodes according to an attribute evolution on $h$. Contrast between nodes is then increased as well as homogeneity inside regions. A delta $\Delta_h$ is set as an evaluation range to apply the MSER for a specific attribute $h$, denoted by $h$-MSER.

Indeed, each $h$ has its own evolution behavior trough node branches which do not necessarily have a correlation with other attributes $h$ (Figure 2.9). For a max-tree $\tau^{max}$, due to their spectral–spatial configuration, area has an incremental behavior whereas intensity level has a decremental evolution. Compacity and barycenter distance are geometric attributes that do not follow a relation order, varying along the branch with their respective inflection points without a defined pattern. If a node $N_i$ is not considered as stable by applying equation (2.2), it is pruned ($\hat{N}$). Following ascendance path, the pixels/voxels of ($\hat{N}$) are placed in the nearest stable node, setting a tree reconfiguration $\hat{\Psi}_h$.

Thus, the initial tree $\Psi$ is submitted to independent and parallel $h$-MSER evaluations to obtain asymmetrically filtered trees $\hat{\Psi}_h$.

### Attribute projection

A pixel/voxel has a multiple node membership, belonging to several nodes from $N_i$ and all its ancestors $N_p(N_i) \subseteq N_{root}$. In the direct restitution criterion, for a scalar image restitution, the highest level $\lambda$ is the most representative for a pixel/voxel, mapping its gray level intensity by:

$$I(\Omega) = \bigvee_{N \in \Theta} C_N, \tag{2.8}$$

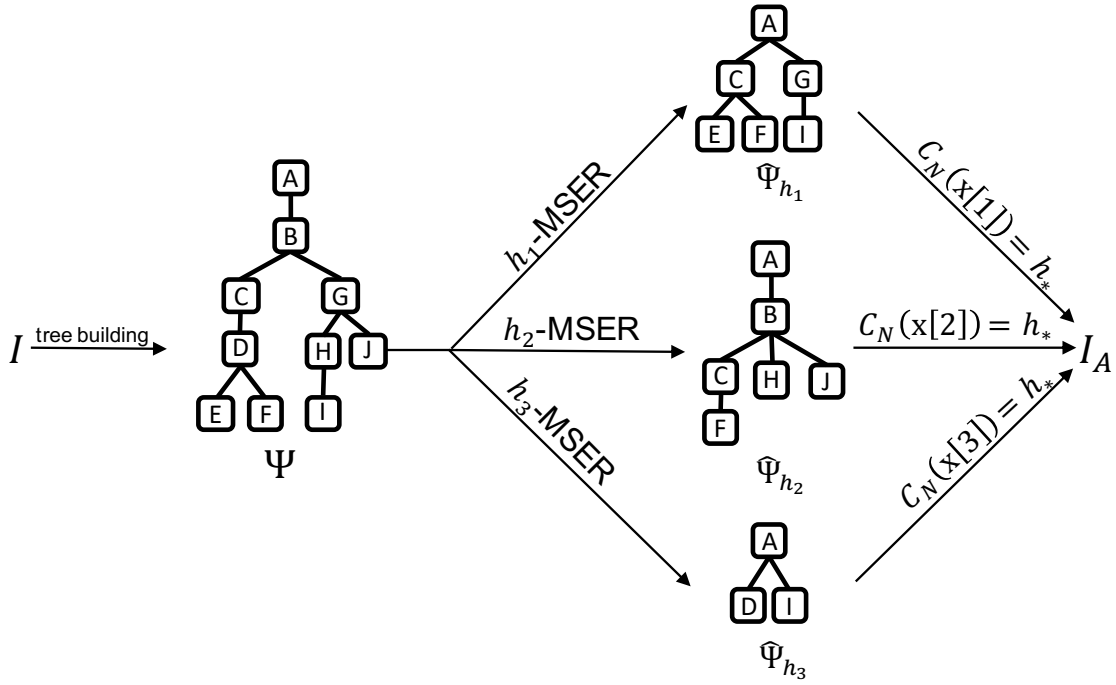


Figure 2.10: Transformation of an input $I$ by a tree processing to build a vectorial image $I_A$. A number $n$ of $h$-MSER filtering for $\hat{\Psi}_{h_n}$ produces a dissimilar family of trees from the original tree $\Psi$ of $I$. This provides the data $h_*$ for $I_A$ building by a cylindrical function $C_N(x[n]) = h_*$.

where $C_N$ is a cylindrical function that allows to map a pixel/voxel $x$ with a node $N$. Thus, $C_N(x[n]) = h_n$ maps an attribute value of a node to a pixel/voxel in a channel $n$ of $I_A(x[n])$.

Each filtered tree $\hat{\Psi}_h$ is subject to a respective tree-image projection $C_N(x[n]) = h_*$ to fulfill a $n$ channel of $I_A(x[n])$, where $h_*$ is a selected node attribute that is extracted from stable nodes $\hat{\Theta}$ of $\hat{\Psi}_{h_n}$ (Figure 2.10).

**Vectorial random walker**

This particular vectorial image $I_A$ with $n$ channels for pixel/voxel $x$, is proposed to be exploited with the conventional vectorial random walker algorithm for image segmentation. This way, our random walker is fed by asymmetric stable region-based data. Weight function in equation (2.3) is adapted to measure vectorial distance between vectorial values $x[h]$ on graph nodes $V$ by:

$$w_{ij} = exp(-\beta[h]\|v_i[h] - v_j[h]\|^2) \tag{2.9}$$

where the norm $\|v_i[h] - v_j[h]\|$ is represented by a Mahalanobis distance $d_m$ which is defined as follows:

$$d_m(v_i, v_j) = \sqrt{(v_i[h] - v_j[h])^T \Sigma^{-1} (v_i[h] - v_j[h])} \tag{2.10}$$

where $\Sigma^{-1}$ corresponds to inverse covariance matrix. This distance type $d_m$ is used to normalize data scattering because we are handling various attributes with different intervals.

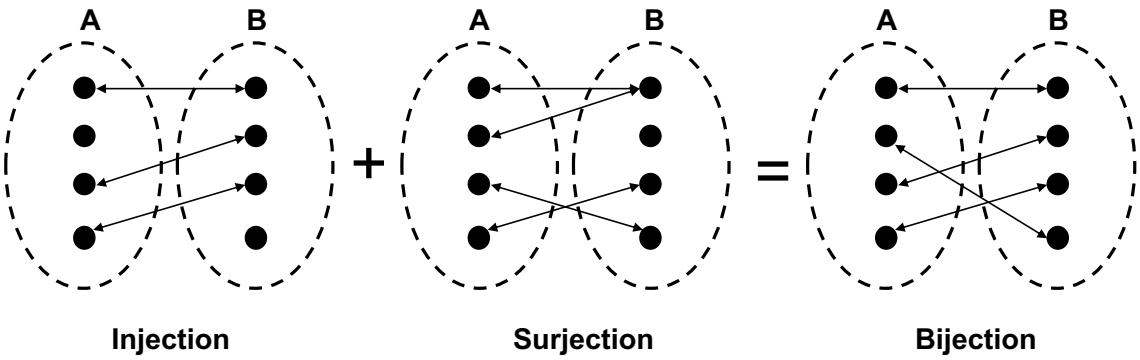


Figure 2.11: Examples of injective, surjective and bijective matching between A and B sets.

According to a region class of interest $c \in C$, pixel/voxel seeds $x_{seeds}$ ($v_{seeds}$) are selected. This can be done interactively by manually selecting pixels/voxels that, with high confidence, are members of a class $c$; or by applying automatic algorithms. For a binary class case, where a foreground class $c$ has relatively high contrast with its background, threshold approaches can be used to simplify the selection of seeds.

Finally, from node seeds $v_{seeds}$ representing $c$ classes, the random walker is applied for segmentation.

**Conclusion** A tree is presented as a performing image structure to improve image data by stability attribute filtering. This improvement is dependent to decided attributes to be analyzed which are injected in a mutichannel image. In addition, multimodal information can be added to increase data richness of the vectorial image in order to feed the random walker algorithm. This algorithm allows to exploit this vectorial image concept to label pixels/voxels although fuzzy contours on the objects. It should be noted that selected input images to be used in this methodology must be spatially correspondent with respect to the object of interest.

### 2.2.2 Hierarchical graph matching filtering for segmentation

In graph theory, graph matching problem focuses on finding a similarity between graph components. If there is an exact graph matching, it is known as a graph isomorphism problem with a bijective relation. Otherwise, it is known as an inexact graph matching, with injective or surjective relations, where not all its components are matched or someones over matched (Figure 2.11).

In this context, our work is focused on finding the best injective matching between two sets and ignoring the components that are not matched. A hierarchical structure is here used for two objectives. Firstly, for image partitioning in order to obtain a collection of nodes (regions) that are individually interpreted as an element of a data set; and secondly, to take advantage of spatial topological inclusion to correct possible false matchings. Region-based attributes are used to characterize nodes and establish a comparison metric to weight potential linking between two node sets.

This matching, taking into account graph structure, aims to respect the hierarchical order between nodes and provide the best possible matching between a set A and a set B.

Four main steps compose this methodology: tree building, computing of attributes, matching of nodes and segmentation.

**1. Tree building**

From an image $I_A$ and an image $I_B$, hierarchical structures $\Psi_A$ and $\Psi_B$ are built. The nature of $\Psi$ depends on its relevance for modeling the objects of interest where image partition convention of $\Psi$ (max-tree, min-tree, tree of shapes,...) must guarantee that the object of interest is defined inside the tree as a node $N$.

**2. Computing of attributes**

Relevant attributes $h$ are computed on the nodes $\Theta$ in order to characterize them. These values, generally scalar values, are normalized and then stored in an attribute vector $A$. Their selection depends on co-representation to describe the object of interest in both data sets (Figure 2.12).

**3. Matching of nodes**

$\Theta_a$ and $\Theta_b$ regions of both trees $\Psi_A$ and $\Psi_B$ are correlated with the purpose of finding the most similar ones. Their similarity level is measured by a distance between node attributes, with a tolerance value $\varepsilon$. Regions which distance is lower than a given $\varepsilon$ are considered as similar and stored as potential candidates for matching. Afterward, mismatching is reduced by a correlation correction step, deleting false positives. This step is explained in detail in Section 2.2.2.1, because of being the core of this approach.

**4. Segmentation protocol**

A protocol of segmentation is applied to identify the nodes belonging to the object of interest. These nodes are re-embedded in their initial image support space to recover the segmentation result. An example of segmentation protocol is presented in Section 2.2.2.2.

#### 2.2.2.1 Matching of nodes

Matching process can be divided in two steps: recovering of candidates and matching selection between candidates.

**Recovering of candidates**

This step aims to find the node candidates for a possible matching.

Nodes $\Theta_A$ of $\Psi_A$ are co-related with the nodes $\Theta_B$ of $\Psi_B$. Thus, the comparison queue for each node $N_a$ is composed by all nodes $\Theta_B$ whereas $\Theta_A$ for $N_b$. Let $M(N_a, N_b)$ be a full surjective mapping function between $\Theta_a$ and $\Theta_b$, with Boolean values. $M(N_a, N_b)$ is used to define the inter-node matching state between two

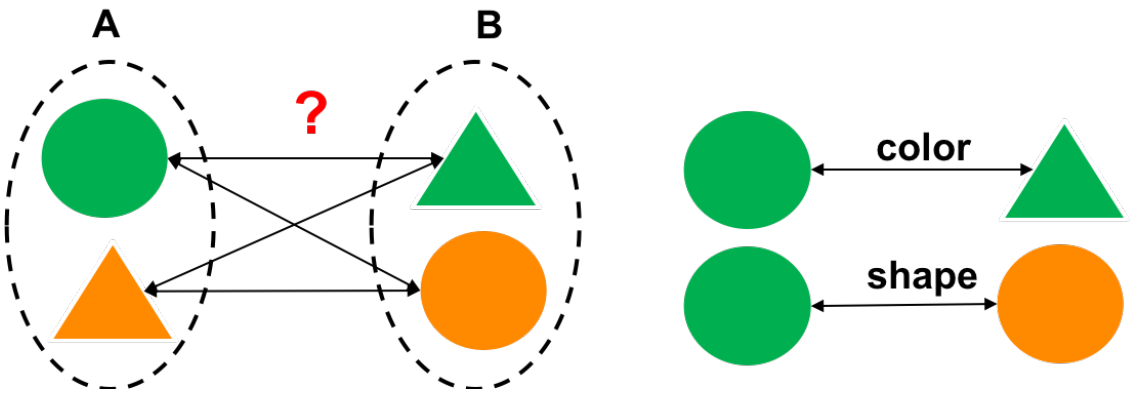


Figure 2.12: Injective matching between A and B sets according to color and shape features.

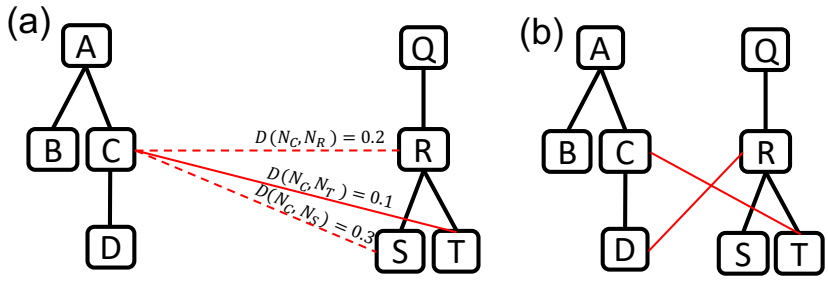


Figure 2.13: Graphical representation of (a) best representative selection and (b) hierarchical crossed problem on two trees $\Psi$. Red lines are the correlation links $M(N_C, [N_R, N_S, N_T])$ where the pointed ones, $M(N_C, N_R)$ and $M(N_S, N_C)$ represent not current best candidate, where $\varsigma(N_C) = N_T$.

nodes $N_a$ and $N_b$, where $M(N_a, N_b) = 1$ refers to an active potential matching and $M(N_a, N_b) = 0$ to not reached matching.

A Euclidean distance $d(\vec{A_a}, \vec{A_b})$ is computed between inter-node attributes $A$:

$$d(\vec{A_a}, \vec{A_b}) = \sum_{i=1}^{H} \beta[i](A_a[i] - A_b[i]), \tag{2.11}$$

and stored as $D(N_a, N_b) = d$, where $\beta$ is a vector of size $H$ that weights up each attribute according to its relevance, with $\sum_{i=1}^{H} \beta[i] = 1$. If $d(\vec{A_a}, \vec{A_b}) < \varepsilon$, a similarity matching $M(N_a, N_b) = 1$ is achieved, otherwise $M(N_a, N_b) = 0$; where $\varepsilon$ is a user-defined parameter chosen within the interval $[0, 1] \in \mathbb{R}$ that establishes the similarity threshold.

If a node $N_a$ or $N_b$ does not have at least one candidate among $\Theta_B$ or $\Theta_A$, expressed by $M(N_a, \Theta_B) == 0$ or $M(\Theta_A, N_b) == 0$, $N_a$ or $N_b$ is considered unmatched since this first step.

#### Matching selection between candidates

A candidate selection phase is here applied considering: the best available candidate and hierarchical crossed problem.

Between possible candidates of $N_a$ (*idem* $N_b$), for $M(N_a, N_b) == 1$, the one that has the lowest value $D(N_a, N_b)$ is set as the best available candidate (Figure 2.13.(a)), denoted by $\varsigma(N_a) = \arg\min D(N_a, \Theta_B)$ or $\varsigma(N_b) = \arg\min D(\Theta_A, N_b)$.

A hierarchical crossed problem occurs when hierarchy at potential matching is not respected (Figure 2.13.(b)). A general formulation can be presented by an ancestor node $N_{anc}(N_a)$ (lower $\lambda*$) related with $N_b$ and an ancestor $N_{anc}(N_b)$ related with $N_a$, or inversely.

Let $\xi(N_a, N_b)$ and $\Xi(N_a, N_b)$ be Boolean functions to establish a current and final matching between $\Theta_A$ and $\Theta_B$, according to a current $sum$ and final $Sum$ distance metric (scalar values).

Following these definitions, an iterative selection process is launched. An iteration is composed of two steps: local ascending and global descending walking.

Local ascending step allows one to evaluate local matching modifications according to candidate election order. The term "ascending" here denotes hierarchical ascending to the root. Thus, $\xi(N_a, N_b)$ matching is conditioned according to a node elector order, giving the privilege of first elector to all the nodes of both trees. The term $N_{a/b}$ is here used to denote indifference for using $N_a$ or $N_b$. A candidate selection thus begins at starting node $N_{start}$ way down to the $N_{root}$, in the branch $\varphi(N_{start}, N_{root})$. In this particular linear path, $N_c(N_{a/b})$ is the child node of $N_{a/b}$ that belongs to the defined branch section $\varphi$. For the current node $N_{a/b}$, thus, its best candidate $\varsigma(N_{a/b})$ is evaluated, to decide if it can be selected for a matching $\xi(N_{a/b}, \varsigma(N_{a/b}))$, denoting a matched node as $\bar{N}_{a/b}$. In the direction to $N_{start}$ on the branch $\varphi(N_{start}, N_{root})$, the first matched child $\bar{N}_c(N_{a/b})$ is searched. If $\varsigma(N_{a/b})$ is not included in the subtree $\psi(\varsigma(\bar{N}_c(N_{a/b})))$ or not matched child $\bar{N}_c(N_{a/b})$ found because of reaching $N_{start}$, then $\xi(N_{a/b}, \varsigma(N_{a/b})) = 1$ and all candidates $M(\Theta_{A/B}, \varsigma(N_{a/b})) = 0$; otherwise $\varsigma(N_{a/b})$ is discarded by $M(N_{a/b}, \varsigma(N_{a/b})) = 0$ and best candidate evaluation is repeated while there are potential candidates.

Global descending walking aims to end tree matching that is conditioned by local ascending matching already established. For a node $N_{a/b}$ not belonging to $\varphi(N_{start}, N_{root})$, its best candidate $\varsigma(N_{a/b})$ is evaluated. In direction to the root $N_{root}$, the first matched parent $\bar{N}_p(N_{a/b})$ is searched. If $\varsigma(\bar{N}_p(N_{a/b}))$ is not included in the subtree $\psi(\varsigma(N_{a/b}))$ or not first matched parent found, $\xi(N_{a/b}, \varsigma(N_{a/b})) = 1$ and all candidates $M(\Theta_A, \varsigma(N_a)) = 0$ or $M(\varsigma(N_b), \Theta_B) = 0$; otherwise $\varsigma(N_{a/b})$ is discarded and remaining candidates are evaluated.

Initializing current weight value $sum = 0$, after each iteration of a $N_{start}$, in coupled nodes $\forall \xi(N_a, N_b) = 1$, a summation in $sum$ is launched $sum = \Sigma exp(-D(N_a, N_b))$. If $sum > Sum$, then $Sum = sum$ and $\Xi = \xi$. The final matching result $\Xi$ thus corresponds to the highest summation weight $Sum$.

#### 2.2.2.2 Object identification for segmentation

On $\Xi$, a segmentation protocol can be applied to select those nodes $N_{a/b}$ that correspond to the object of interest. The fact of selecting a $N_a$ or $N_b$ means to parallely point to a node $N_b$ or $N_a$ for co-segmentation. An example of segmentation protocol can be designed by partially detecting some pixels/voxels for the foreground $fg$ on the image $I$, denoting this pixel/voxel set by $x_{fg}$, to couple with their respective nodes $N_{fg}$ in the tree $\Psi$. For example, a threshold-based pre-segmentation, by a percentage $T_{fg} \in \mathbb{R}[0, 1]$ on $I(q_{max})$ can be used to select pixels/voxels $x$ which intensity $q$ respects this partition condition:

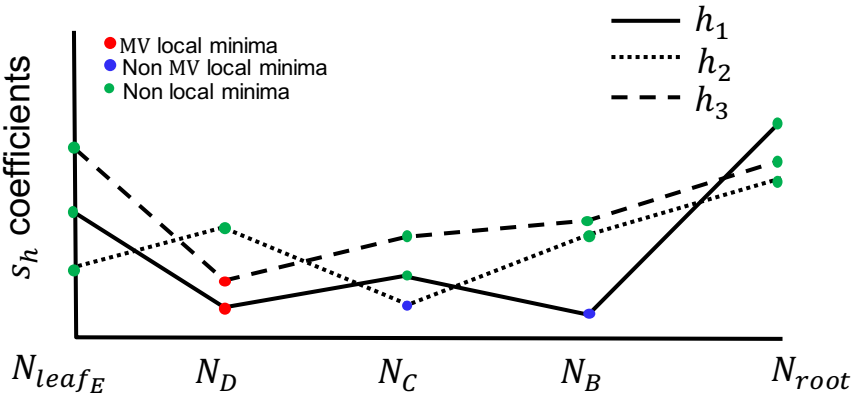


Figure 2.14: Example of majority vote (MV)-MSER signal analysis on the branch $\varphi_1(N_E, N_A)$ for arbitrary attributes $h_1$, $h_2$ and $h_3$.

$$\Omega_{fg} = I(\Omega) > q_{max} * T_{fg}. \tag{2.12}$$

where $\Omega_{fg}$ is the support of pixels/voxels that are marked in foreground class. Actually, this $N_{fg}$ node detection has a semantically different meaning compared with a $x_{fg}$, because pixel/voxels seeds can be placed in a node where other ones do not, due to tree partitioning type and matching filtering.

Finally, $\Psi_A$ and $\Psi_B$ are reconfigured by pruning non-selected nodes $N_{a/b}$ as $N_{fg}$, replacing their pixels/voxels to their nearest active parent $N_{fg}$. Both trees are back-projected to obtain the matrix image structure, by $C_{N_{fg}}(x) = 1$, those that reached root level by $C_{N_{root}}(x) = 0$, producing a binary segmentation.

**Conclusion** This methodology aims to match similar objects from two images ignoring the spatial correspondence between images or the objects of interest, preserving a hierarchical coherence. If these images have a semantic difference because of different mode source (multimodality), this approach can help to remove not relevant information, preserving similar one. Thus, objects which are not defined in both trees, according to respective partition protocol, are filtered, and those ones which are preserved can be detected for segmentation. The quality of these matched objects is validated by the co-representation under selected attributes. Indeed, the result is high dependent of discrimination level of selected attributes which is used to define the current similarity concept.

### 2.2.3 Majority vote MSER filtering and vectorial random walker on hierarchical graph

Based on two previous contributions, we propose an extension of the tree-based segmentation process. This methodology explores the whole tree taking advantage of its structural characteristics to simplify regional image interpretation, filter unstable regions in order to apply a random walker segmentation on the tree.

Three steps are the main core of this method: tree computing, majority vote MSER filtering and random walker segmentation.

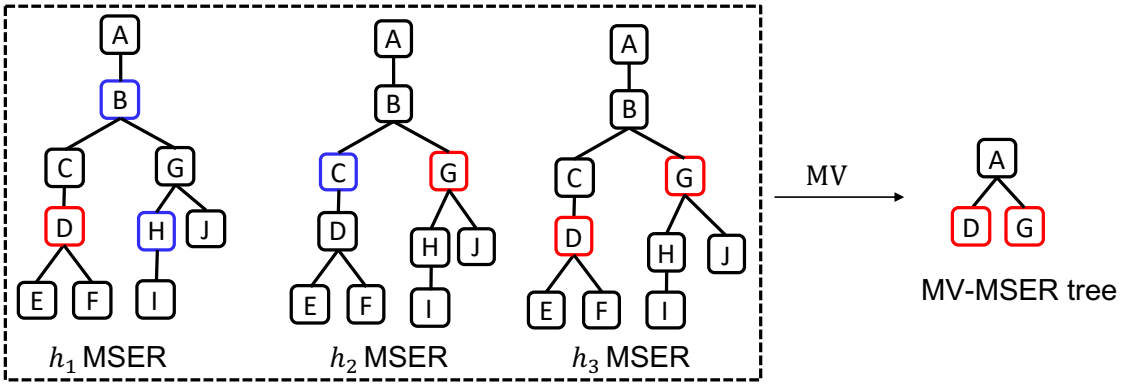


Figure 2.15: Graphical representation of majority vote (MV)-MSER with colored boundary voters (red for majority and blue for minority). Decreasing node votes are for $N_D = 2$, $N_G = 2$, $N_B = 1$, $N_C = 1$, $N_H = 1$ and remaining ones are 0. By majority vote, just $N_D$ and $N_G$ are stable nodes (red), unstable nodes are pruned excepting the root as support of the tree.

### Majority vote MSER

MSER stability metric on hierarchical structures was presented to analyze region-based attributes and preserve a stable tree configuration. A number $n$ of region-based attributes can be co-evaluated to design a more robust stability metric. Indeed, a local minimum definition can be disturbed by local noise such as a non-representative local minimum. A stability agreement, by majority vote statement is evaluated to support stability from several region-based MSER. From $h_1$ to $h_n$, if most of $s_h(N_i)$ are local minima, then $N_i$ is considered as a majority vote stable node (Figure 2.14). Unstable nodes are then discarded from the tree, and their pixels/voxels are shifted to the nearest stable parent, re-configuring tree structure (Figure 2.15).

### Random walker on a tree

Random walker approach can be applied to any graph structure to label nodes based on the shortest path algorithm. Thus, it can be adapted for a hierarchical graph configuration $\Psi(N, E)$ with multi pixel/voxel nodes $N \in \Theta$ and child/parent hierarchical edges $e \in E$. However, there are some structural notions on a tree that can be exploited to optimize and redefine the random walker algorithm. With no need of computing, descendants of a node seed, if there are not other seed class in their subtree $\psi$, can be directly assigned to their respective seed class. Moreover, tree hierarchical configuration can also be used to directly analyze the spatial evolution of a region by evaluating its direct ancestor line in the branch.

Let a transition zone, as a specific type of a branch section $\varphi$, denoted by $\kappa(N_{seed}(c_1), N_{seed}(c_2))$, be an uncertainty space of unmarked nodes $N_U \in \Theta_U$ between two seed nodes $N_{seed}(c_1)$, $N_{seed}(c_2) \in N_{seeds}$ from two different classes $c_1$ and $c_2$.

A random walker can be applied on $\kappa(N_{seed}(c_1), N_{seed}(c_2))$ in order to assign to $N_U$ their respective class $c_1$ or $c_2$ according to shortest path paradigm between unmarked nodes $\Theta_U$ and respective class node seeds $N_{seeds}$.

A node $N_U$ can belong to $n$ uncertain branches $\kappa$, thus $n$ possible labeling as $c_1$ or $c_2$. This label overlapping as well as the final segmentation, can be solved by defining how this labeling is interpreted in the tree.

The hierarchical arrangement of nodes is projected to the establishment of two types of node class: passive and active node classes.

A passive node class $c_*^-$, generally used for the background, provides to their node members the particularity of a negative hierarchy. Thus, if there is no active node seeds, a node $N_i$ and all its ascendant nodes in its branch line $\varphi(N_i, N_{root})$ are labeled as imperative members of class $c_*^-$, denoted as $N_i(c_*^-)$.

For an active node class $c_*^+$, all the descendents of a node $N_i$ (a subtree with root $N_i$ by $\psi(N_i)$) are imperative members of class $c_*^+$ with no exception, neither a passive class node $N_i(c_*^-)$ nor other positive class node $N_i(c_*^+)$.

Thus, considering a hierarchical order where the root $N_{root}$ is the the highest level and the leaves $N_{leaves}$ are the relatively lowest level; the supremum $N_i(c_*^+)$ at $\Psi$ is the segmented object for each positive class $c_*^+$.

Seeds can be selected in the image $I$ and then mapped to $\Psi$ to find seed nodes (as in Section 2.2.2.2 by thresholding), or they can be directly selected on the tree $\Psi$ by providing node tracks.

Defining the number and type of region-based attributes according to object of interest characteristics, the level of similarity between a node $N_i$ and $N_j$ can be defined by a weight function:

$$w_{ij}(e_{ij}) = 1 - \sum_{h=1}^{H} \beta[h]|A_i[h] - A_j[h]|, \tag{2.13}$$

where $\beta$ is a vector of size $H$ that weights up each attribute according to its relevance, with $\sum_{h=1}^{H} \beta[h] = 1$.

**Algorithm**

In line with the qualities of the object of interest, a selected type of tree is built from the input image $I$, obtaining a tree $\Psi$ of $I$. From the set nodes $\Theta$, a $H$ number of considered relevant region-based attributes $A$ are computed. Majority vote MSER from $H$ attributes stability is applied to filter the tree $\Psi$ rendering a pruned tree $\hat{\Psi}$, preserving its stable attribute regions in a consensus.

A number $n$ of classes $c$ are here defined where there is at least one positive class (object or foreground) $c_*^+$ and only one background class $c_{bg}^-$.

The protocol selected for the definition of node seeds is then applied in the tree $\hat{\Psi}$, providing with at least one node $N_i(c)$ to each class $c$ where the $N_{root}$ is enhanced as member of background class $c_{bg}^-$.

Hierarchical class properties, negative for $c^-$ and positive for $c^+$, are then interpreted. Propagating ascending and descending class labels, a bipolarization of node branches $\varphi(N_{leaf}, N_{root})$ is obtained, wherein at most one uncertain branch zone $\kappa(N_i(c_*^+), N_i(c_{bg}^-))$ is found. Indeed, potentially in this bipolar propagation, unmarked nodes $N_U$ are labeled for all $c^+$ and $c^-$ classes; specially for positive propagation where nested nodes $N_j(c_{bg}^-)$ in $N_i(c_*^+)$ are re-labeled to class $c_*^+$.

Through the tree nodes, uncertain sections $\kappa(N_i(c_*^+), N_i(c_{bg}^-))$ are searched to trigger the random walker algorithm, fed by vectorial attributes $A$, and thus com-

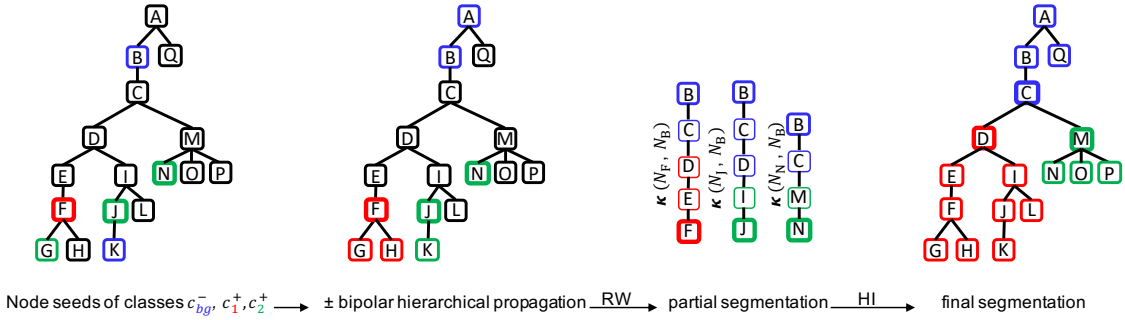


Figure 2.16: Algorithm sequence for random walker segmentation (RW) on a tree $\Psi$. Node seeds $N_{seeds}$ for classes $c_{bg}^-$ in blue, $c_1^+$ in red and $c_2^+$ in green. HI refers to hierarchical interpretation to obtain the final segmentation result. Thickened box node borders represent the supremum node $N$ of class $c_*^+$ at local branch $\varphi$, whereas for class $c_{bg}^-$, the infimum node $N$.

pleting nodes labeling. The obtained probability map $p_U$ (equation (2.7)) is thus interpreted by: if $p_U(c_*) \geq 0.5 = 1$, otherwise $p_U(c_*) = 0$, binarizing the result.

A node $N$ can belong to a number $n$ of branches $\kappa$ in the tree $\Psi$ thus, $n$ segmentation results for membership assignments to a class $c_*$. By a hierarchical interpretation, the node $N_{c_*^+}$ that reaches the highest hierarchical level, has the privilege for node labeling. If a branch $\varphi$ is just composed of nodes belonging to the class $c_{bg}^-$, thus its nodes are labeled at this passive class. The final segmentation is thus composed by the subtree $\psi$ of respective supremum nodes $N(c_*^+)$.

An example of random walker on a tree $\Psi$ is shown in Figure 2.16. In bipolarization step, active class properties provokes the $N_{\mathrm{G}}, N_{\mathrm{K}}$ relabeling by higher hierarchy of nodes $N_{\mathrm{F}}$ and $N_{\mathrm{J}}$, respectively, which, together with $N_{\mathrm{M}}$, are supremum local nodes of object classes $c_*^+$. These latter are here elements, jointed with node $N_{\mathrm{B}}$ of class $c_{bg}^-$, form uncertain branches $\kappa$, where the random walker is applied. In the hierarchical interpretation step, after random walker result, the node $N_{\mathrm{Q}}$ is assigned to $c_{bg}^-$ because of belonging to an isolated branch that was not intercepted in its ascending path by positive class; subtree $N_{\mathrm{I}}$ firstly belonging to $c_2^+$ is relabeled to class $c_1^+$ because its ascendant node $N_{\mathrm{D}}$ reached that label class. Finally, supremum nodes $N_{\mathrm{D}}$ and $N_{\mathrm{M}}$, including respective subtrees $\psi$, represent the segmented objects.

**Conclusion** This methodology shows its advantages by exploiting the tree structure. Firstly it allows to spatially and spectrally organize the image data as nested regions. Secondly, it uses this configuration for automatically deduce regions memberships along the algorithm steps, reducing time computing. Majority vote MSER is a higher discriminant stability metric (used for filtering) based on the consensus of individual attribute MSER analysis. The random walker applied on the tree has the meaning of independently node evolution with the notion of how much distance (number of nodes) a walker (from a node start) can go through the weight that this path costs to reach either a foreground or the background node. The region that reaches a higher evolution with respect to the root, is going to define the labeling fo all its descendants.

The data used for the definition of region-based attributes can be directly re-

covered from the image-tree or added from other image sources (multimodality), by spatial correspondence, as contextual information enriching attribute definition. Thus, image base, taking advantage of high object contrast, allows to define the object of interest inside a tree structure to be processed, whereas the contextual images, is used as a pertinent complement to increase regional information and thus support the segmentation process. In addition, if contextual information is not relevant to define the object of interest (such as low contrast tumor-to-background), its negative impact on the result is low because of the support on the other attributes.

# Chapitre 3 : Instanciation méthodologique en TEP / TDM

Dans ce chapitre, les trois méthodes décrites dans le chapitre 2 sont expérimentées sur des images médicales en TEP et TDM à des fins de segmentation tumorale. Cette problématique est liée à la difficulté de définir les contours des tumeurs en TEP à cause de sa faible résolution spatiale et en TDM à cause du faible contraste dans les tissus mous dont les tumeurs. Cette multimodalité peut fournir des informations très riches pour l'amélioration de la segmentation. Le terme voxels va être ici utilisé car ce travail porte sur des images 3D. Les caractéristiques générales de notre base de données sont présentées dans la section suivante. Une deuxième section porte sur les spécificités de chaque méthode dans sa mise en œuvre et permet de présenter les expériences et les résultats.

## Paramètres généraux

Au total, nous disposons de 33 séries d'images de patients dont 12 cancers du poumon, 8 cancers de la tête et du cou, 3 cancers du sein, 3 lymphomes, 2 cancers du foie, 3 cancers du col de l'utérus, 1 cancer du côlon et 1 sarcome ; où 25 sont des images TEP / TDM et 8 des images monodate TEP / IC (injection de contraste)-TDM. En outre, 6 patients atteints d'un cancer de la tête et du cou diposent d'images multitemporelles IC-TDM dues au suivi du patient. Une description plus détaillée de notre base de données est présentée dans l'annexe B.

### Sélection des structures hiérarchiques pour la TEP et la TDM

En TEP, où une tumeur est généralement associée à des intensités élevées, la modélisation tumorale en tant que nœud $N_{\mathrm{TEP}_{tumeur}}$ est limitée par les restrictions liées à l'image : faible résolution spatiale entraînant un effet de volume partiel et bruit physiologique lié à la position anatomique. Une tumeur peut être représentée partiellement, ceci étant dû à la proximité d'autres zones de haute intensité. Si l'expert médical souhaite segmenter une zone tumorale active, un arbre de coupes $\tau_{TEP}^{max}$ est une meilleure option car il permet d'ignorer les zones intra-tumorales de faible intensité. Si on souhaite repérer des nécroses ou des cavités à l'intérieur d'une tumeur active, notamment si une radiothérapie est envisagée, il est en revanche préférable de choisir l'arbre des formes pour définir des régions par isocountours en prenant en compte des concavités spectrales à dire les zones de moindre activité que leur contour dans les tumeurs.

En TDM, une tumeur présente souvent un faible contraste par rapport aux

tissus mous voisins, ce qui peut induire une définition floue des frontières de la tumeur, et ne permet pas sa définition de région (nœud) $N_{\mathrm{TDM}_{tumeur}}$ dans l'arbre. Un agent de contraste peut être injecté par voie intraveineuse au patient pour une IC-TDM, dans le but d'améliorer la définition des contours de la tumeur, par le biais de sa vascularisation. Au demeurant, les tissus mous ne présentent pas les intensités les plus élevées ou les plus basses dans un scanner, où les os ont la densité la plus forte pour un tissu biologique, donc le niveau d'intensité le plus élevé (> 400 unités de Hounsfield (HU)) ; tandis que l'air est le plus bas ($\sim -1000$ HU). Donc un seuil osseux $\mathrm{TDM}_{SKF} = \mathrm{TDM}\,(\Omega(q)) < 400$ HU est approprié pour enlever le tissu osseux et améliorer la partition. En général, l'arbre de formes est le type d'arbre qui, au regard des inconvénients de la TDM, peut permettre une représentation plus robuste pour les modalités suivantes : TDM, IC-TDM, $\mathrm{TDM}_{SKF}$ et $\mathrm{IC\text{-}TDM}_{SKF}$.

### Sélection des graines voxel et nœud

Les voxels considerés comme graines, notés $x_{graines}$, peuvent être définis intuitivement par une sélection manuelle, interactive ou de manière automatique. Dans cette expérimentation, nous définissons deux types de classes : l'objet ($fg$), ici une tumeur, et le fond ($bg$). Les tumeurs sur la TEP ont pour caractéristique principale d'être représentées par des valeurs de haute intensité due à une fixation élevée des biomarqueurs. Un seuil $T_{fg}$ peut être sélectionné pour obtenir des voxels tumoraux partiels (graines). Un seuil $T_{bg}$ pour des zones non tumorales peut aussi être défini afin d'initialiser l'algorithme de segmentation.

### Métrique de validation

La validation des résultats expérimentaux est effectuée en utilisant le coefficient de similarité de Dice (DSC) comme métrique de comparaison pour mesurer le recouvrement spatial (faux positifs et faux négatifs) comme suit :

$$\mathrm{DSC}(V_{gt}, V_{seg}) = \frac{2|V_{gt} \cap V_{seg}|}{|V_{gt}| + |V_{seg}|}, \tag{2.14}$$

où $V_{seg}$ est le volume segmenté et $V_{gt}$ sa vérité terrain. Notre vérité terrain est fournie par un expert de l'Institut Jean Godinot. Celle-ci est obtenue par segmentation manuelle effectuée sur des images TEP / TDM fusionnées.

## Méthode 1

Un marcheur aléatoire guidé par des attributs basés sur les régions est appliqué ici pour la segmentation TEP / IC-TDM dans un contexte multimodal. Ces images médicales assurent une correspondance spatiale et une représentation précise de la tumeur à segmenter. Les étapes et les détails de ce traitement sont présentés ci-dessous.

### Experimentation

Dans cette expérimentation, les images TEP et IC-TDM$_{SKF}$ sont utilisées comme entrée pour les arbres potentiels : $\tau^{max}_{\text{TEP}}$, $\phi_{\text{TEP}}$ et $\phi_{\text{IC-TDM}}$ afin de modéliser les combinaisons d'arbres respectives.

La sélection des attributs est également définie pour TEP et IC-TDM comme suit : $A_{\text{TEP}} = A_{\text{IC-TDM}}$ = [aire, coordonnées du barycentre, compacité , complexité, Scont]. Cette sélection d'attributs permet de considérer de manière uniforme les sources de données et l'extraction d'attributs standard à partir des champs d'application : spatial (aire et coordonnées du barycentre), géométrique (compacité et complexité), spectrale (contraste) et hybride (Scont (contraste spatial)). Une évaluation avec un $\Delta = 5$ est effectué de manière homogène pour un nombre $n$ de $A[h]$-MSER appliqués aux arbres $\Psi_{\text{TEP}}$ et $\Psi_{\text{IC-TDM}}$ afin de réaliser une analyse hétérogène des attributs et ainsi obtenir $n$ arbres filtrés $\hat{\Psi}_h$ à chaque arbre $\Psi_{\text{TEP}}$ ou $\Psi_{\text{IC-TDM}}$. Dans la reconstruction vectorielle $I_A$, la pertinence de la projection des attributs basée sur la région est évaluée pour $h_* = \lambda_q$, $I_{A_q}$ et pour $h_* = h$. La multimodalité est représentée sur l'image vectorielle $I_A$ en concaténant, au canal des attributs, la matrice vectorielle individuelle TEP$_A$ et IC-TDM$_A$ pour 8 canaux $I_A(x) = [x_{\text{TEP}}, x_{\text{IC-TDM}}]$.

Les résultats ont été comparés entre les versions vectorielle et scalaire pour le marcheur aléatoire a été appliqué aux images monomodales : TEP et IC-TDM ; et multimodales : TEP / IC-TDM.

### Conclusion

Les résultats du marcheur aléatoire aussi bien sur des données scalaires que vectorielles, montrent que la multimodalité augmente la précision des résultats. Ceci corrobore les résultats précédemment obtenus dans la littérature. En effet, leur complémentarité fonctionnelle et morphologique permet une meilleure segmentation des tumeurs. Pour les tumeurs pulmonaires, ce fait est plus évident en raison du contraste partiel élevé entre les tissus mous et la cavité thoracique.

Le marcheur aléatoire vectoriel, représenté par les canaux contenant une information issue des arbres filtrés par $h$-MSER, affiche une performance supérieure à celle du scalaire. En effet, la stabilité des attributs dans les structures hiérarchiques nous permet de mieux filtrer le bruit dans l'image. La polyvalence des attributs et de la sélection des modalités d'imagerie permet d'adapter l'algorithme en fonction des propriétés inhérentes à la cible. De plus, les attributs peuvent aider à différencier des régions qui, uniquement en intensité, sont considérées comme similaires.

## Méthode 2

Le filtrage par correspondance de graphes hiérarchiques est expérimenté ici sur un ensemble de données multitemporelles pour la segmentation de tumeurs. On considère le paradigme de couplage fonctionnel / anatomique afin de trouver des régions (telles que des tumeurs), présentes dans les deux images TEP / IC-TDM, via le guidage par des attributs basés sur les régions. Par conséquent, la sélection

des attributs est très importante à des fins de discrimination. Le cerveau est choisi comme référence spatiale importante lorsque aucune correspondance existe.

## Détection du cerveau sur TEP et TDM pour la normalisation co-spatiale

Le cerveau est un organe qui peut être automatiquement détecté sur les deux modalités d'imagerie et exploité comme référence anatomique précieuse pour le couplage entre la TEP et la TDM.

En TEP acquise par injection de fluorodésoxyglucose (FDG), le cerveau est une zone à fixation élevée. Cette propriété peut être exploitée pour segmenter plus facilement le cerveau en raison du contraste relativement élevé avec l'arrière-plan. Un seuillage par un pourcentage $T_{cerveau}$ de la TEP($q_{max}$), calculé comme suit, $\text{TEP}_{bw} = \text{TEP}(\Omega) > T_{cerveau} * \text{TEP}(q_{max})$, peut être appliqué sur une image TEP pour obtenir une segmentation binaire $\text{TEP}_{bw}$. Cependant, la fixation dans le cerveau n'est pas souvent homogène, y compris dans les cavités crâniennes non actives, notamment les ventricules remplis de liquide céphalorachidien. Ce problème peut être résolu en appliquant une opération de fermeture pour fusionner les zones correspondant au cerveau en un seule objet. Une fermeture $\bullet$ est l'érosion $\ominus$ de la dilatation $\oplus$ d'un ensemble $A$ par un élément structurant $SE$ par : $A \bullet SE = (A \oplus SE) \ominus SE$ ; où les objets $o$ (composantes connexes) de support $O$, selon une séparation spatiale $SE$ de taille et forme inférieure entre eux, sont fusionnés. Soit $A = \text{TEP}_{bw}$ et $SE$ une sphère de 2 cm de rayon, entre les objets binaires obtenus par $\text{TEP}_{bw} \bullet SE$, l'objet $o_i$ situé en position anatomique la plus élevé suivant l'axe z (plan axial) avec la plus grande taille, est considéré comme $o_{cerveau_{\text{TEP}}}$.

En tomodensitométrie (TDM), où le cerveau est composé de matière blanche et grise (20–45 HU) et de cavités intracrâniennes, on observe un contraste périphérique élevé avec le crâne mais une différenciation faible (pratiquement nulle) avec la moelle épinière (40–50 HU) et les nerfs périphériques. Un filtrage osseux (*skeleton filter-SKF* en anglais)$\text{TDM}_{SKF} = \text{TDM}(\Omega(q)) < 400$ HU est efficace pour supprimer le crâne et la colonne vertébrale en tant qu'enveloppe principale du système nerveux central. Les connexions cerveau / moelle épinière / nerfs périphériques restantes peuvent être rompues par une opération d'ouverture $\circ$, consistant en la dilatation $\oplus$ de l'érosion $\ominus$ d'un ensemble $A$ par un élément structurant $SE$ : $A \circ SE = (A \ominus SE) \oplus SE$, où $A = \text{TDM}_{SKF}$ et $SE$ est une sphère de 7 *mm* de rayon en tenant compte que la moelle épinière présente une épaisseur moyenne de 13 *mm* au niveau cervical. En appliquant un seuil sur $\text{TDM}_{SKF} \circ SE > 20$ HU pour la binarisation d'images, l'objet $o_{\text{TDM}_{cerveau}}$ est le $o_i$ connecté positionné dans le niveau axial le plus élevé avec la plus grande taille.

## Expérimentation

L'ensemble multitemporel, composé de patients atteints de cancer du cou et de la tête, est utilisé ici à des fins d'expérimentation. Cette région anatomique est caractérisée par une forte présence de structures osseuses telles que le crâne, les vertèbres cervicales, qui peuvent perturber la définition de la tumeur. Un filtrage du squelette est appliqué dans la IC-TDM pour obtenir une $\text{IC-TDM}_{SKF}$.

Un arbre de formes $\phi$ est construit à partir de chacune des images TEP et IC-TDM. Les attributs basés sur des régions sélectionnés doivent être représentatifs des tumeurs pour les deux modalités d'image dans un cadre global. L'aire multipliée par l'unité d'espacement d'image respective (la taille du voxel 3D) constitue une métrique spatiale comparable. Il en est de même pour la compacité en tant que région géométrique adimensionnelle ; et finalement, pour le barycentre $\mathrm{BC}_{bn}$ normalisé en soustrayant le barycentre du cerveau.

Une fois l'arbre construit et les attributs calculés, des filtres sont appliqués pour améliorer la qualité des données en supprimant les informations bruitées ou non pertinentes. Nous profitons des caractéristiques de traitement des arbres anti-extensifs pour ne pas ajouter de données supplémentaires, mais simplement pour supprimer les nœuds relativement petits inférieurs à $1cm^3$ par filtrage d'aire, considérés ici comme inférieurs aux tumeurs traitées.

Après la procédure de couplage, la TEP (seuil $T_{fg} = 0,6$) est utilisée pour détecter les nœuds $N_{\mathrm{TEP}_{tumeur}}$ dans l'arbre $\hat{\phi}_{\mathrm{TEP}}$ en raison de sa haute spécificité. Le nœud $N_{\text{IC-TDM}_{tumeur}}$ est obtenu par la relation de couplage $\Xi(N_{\mathrm{TEP}_{tumeur}}, N_{\text{IC-TDM}_{tumeur}})$, avec le nœud $N_{\mathrm{TEP}_{tumeur}}$.

Pour la validation, les versions monotemporelles (seulement sur la TEP) du marcheur aléatoire sont appliquées. Pour le marcheur aléatoire scalaire, nous considérons les paramètres initiaux suivants : $\alpha_1$ et $\alpha_2$ sont définis sur 0,5 pour considérer également les probabilités $p$ de chaque modalité dans leur fusion et, $T_{fg} = 0,6$ et $T_{bg} = 0,15$ qui sont des seuillages standard définies expérimentalement. Ces mêmes paramètres initiaux sont aussi utilisés pour le marcheur aléatoire vectoriel, ainsi que des paramètres supplémentaires spécifiques : le vecteur d'attributs $A_{\mathrm{TEP}}$ = [Area, BC, compacité, complexité, contrast, SCont], $\Delta = 5$ et les graines propagées par les arbres $\tilde{v}_{graines}$ par $\hat{\phi}_{TEP}$. Ce paramétrage correspond à celui proposé dans la méthode 1.

## Conclusion

La mise en correspondance spatiale entre les images médicales est généralement problèmatique, en particulier lorsque les informations TEP et IC-TDM issues d'une acquisition multitemporelle sont précieuses pour le traitement de la tumeur, et qu'un recalage d'image est considéré comme chronophage et n'est pas parfait en raison de son considérable risque d'erreur.

L'arbre de formes semble constituer une solution propice à l'évaluation pour des objets qui ne se correspondent pas dans l'espace. Une fois construite, cette structure peut être interprétée comme étant spatialement indépendante dans le paradigme nœud à nœud. En d'autres termes, les objets d'intérêt peuvent être répartis dans l'arbre, et ce indépendamment de la localisation des objets dans l'image.

Les résultats montrent une bonne précision spatiale, supérieure à celle des méthodes à seuil. Il convient de noter que les tumeurs sont couplées de manière efficace. La normalisation du barycentre cérébral est une étape importante pour améliorer l'analyse spatiale, tirant parti de la co-représentation cérébrale, ainsi que des attributs basés sur les régions sélectionnées, en facilitant la détection de la tumeur à la fois en TEP et en TDM.

# Méthode 3

Dans cette section, le filtrage par vote majoritaire MSER et le marcheur aléatoire sur un graphe hiérarchique sont expérimentés sur des images TEP, TDM et IC-TDM. Ceci permet de prendre en compte les propriétés des images soit dans la phase de définition des régions, soit directement dans la phase de segmentation y apportant de l'information contextuelle. Ces deux rôles possibles doivent être sélectionnés en fonction de leur impact sur les résultats.

## Expérimentation

Des sources monomodales et multimodales sont analysées dans des expériences locales de la méthode. Les images TEP, TDM et IC-TDM sont les entrées potentielles à utiliser dans deux rôles possibles : $I_{\text{base}}$ pour la construction de l'arbre et $I_{\text{contexte}}$ en tant qu'informations complémentaires dans l'arborescence, noté $\Psi(I_{\text{base}}, I_{\text{contexte}})$. Le symbole – est appliqué pour représenter une non-utilisation en tant que $I_{\text{contexte}}$. A partir de $I_{\text{contexte}}$, un nouveau type d'attribut basé sur une région est présenté : attribut d'information de contexte. Les attributs basés sur les régions sont sélectionnés en fonction du rôle de l'image dans l'arbre, créant ainsi un vecteur $A = [A_{\text{base}}, A_{\text{contexte}}]$.

Un filtrage par vote majoritaire MSER est ensuite appliqué sur $\Psi$ avec un delta $\Delta = 1$. Deux classes sont définies ici : les zones tumorales en tant que classe positive $c^{+}_{fg}$ et le fond en tant que classe négative $c^{-}_{bg}$.

Pour la segmentation par marche aléatoire sur l'arbre $\Psi$, $T_{fg}$ et $T_{bg}$ sont définies par les valeurs 0,6 et 0,15, tandis que $\beta[h]$ est fixée de manière homogène à $1/H$ où $H$ est le nombre d'attributs.

Notre ensemble de données a également été expérimenté avec : un seuil classique de 40% sur $I(q_{max})$ ; les *fuzzy-c-means* (classes de la tumeur et du fond) ; et le marcheur aléatoire TEP / TDM avec $x_{fg}$ et $x_{bg}$ comme graines $v_{graines}$ ; où notre méthode sur $\phi$(TEP, TDM) reste la plus performante.

## Conclusion

La tomodensitométrie est peu contrastée pour la tumeur par rapport au fond, ce qui est souvent insuffisant pour établir une représentation régionale par un arbre de formes. Cependant, les données de tomodensitométrie prises comme une information contextuelle modélisée sur un arbre TEP peuvent être une caractéristique importante pour guider la segmentation de la tumeur. En effet, les tissus mous sont difficiles à identifier entre eux ; cependant, leur contact partiel avec les os et les cavités remplies d'air peut fournir des informations précieuses pour détecter les frontières partielles.

La TEP modélisée sous la forme d'un arbre de formes contenant des informations contextuelles de la tomodensitométrie, constitue l'expérimentation moyenne la plus précise. Ces résultats confirment le fait que la multimodalité peut améliorer les performances de la segmentation. Ce fait est notamment cohérent, puisque la vérité terrain est définie sur un périmètre bimodal fusionné et pas uniquement sur la TEP.

Notre méthodologie repose sur une forte interaction entre ses composants principaux afin de favoriser sa robustesse : un arbre en tant que structure d'image pour le filtrage par vote majoritaire MSER et pour la segmentation par marcheur aléatoire. L'arbre de coupes et l'arbre des formes sont des structures pertinentes pour traiter les régions d'image à haute intensité, avec une sémantique tumorale. Le vote majoritaire MSER présente des avantages certains pour lisser les frontières floues et pour ignorer les nœuds non-importants comme le bruit et obtenir ainsi une structure réduite à 2% de ses nœuds d'origine. Le marcheur aléatoire, en tant que méthode de segmentation, est uniquement appliqué aux sections de branche d'incertitude (pas à tout le graphe) en tirant parti des qualités structurelles des arbres. Ces couplages méthodologiques, basés sur les arguments mentionnés ci-dessus et les résultats obtenus lors des expérimentations cliniques, semblent être robustes face à la segmentation tumorale par TEP / TDM.

# Chapter 3

# Methodological instantiation in PET/CT

## Contents



In this chapter, methodologies described in Section 2.2.1: Method 1, in Section 2.2.2: Mehod 2 and in Section 2.2.3: Method 3, are experimented on PET and CT medical images for tumor segmentation purpose. Voxels term is going to be used as image unit element $x \in \Omega$ because of using 3D images. General characteristics of our input data set are introduced in the next Section. In a second Section, we present each method specificities for implementation as well as experiments and results.

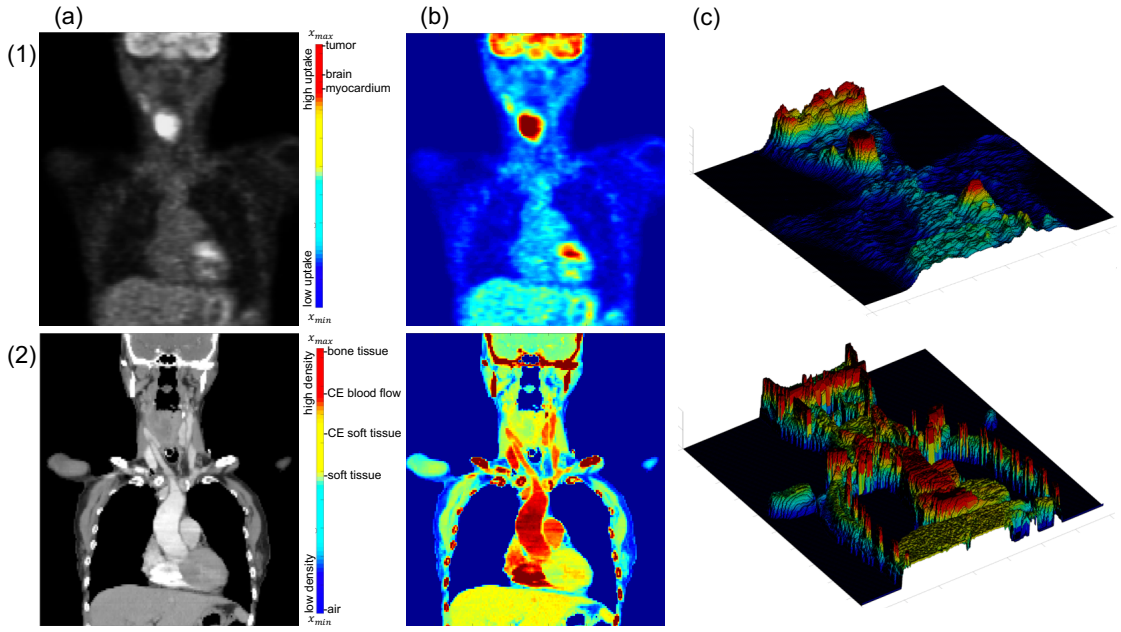


Figure 3.1: PET/CT intensity interpretation. (a) Gray level, (b) colormap and (c) 3D relief of 2D sagittal slice from (1) PET and (2) CE-CT.

## 3.1 General parameters

Our medical data set is classified according to medical imaging type, cancer type and temporal frame, referring to a sequential PET/CT acquisition from a unique scanner, known as monotemporal framework, or a CT alone in a different device for a multitemporal one. PET/CT hierarchical structure setting is analyzed and explained. A list of potential PET/CT region-based attributes is presented.

In total, we have 33 patient image sets with 12 lung cancer, 8 head an neck cancer, 3 breast cancer, 3 lymphoma, 2 hepatic cancer, 3 cervix cancer, 1 colon cancer and 1 sarcoma; where 25 sets are PET/CT and 8 PET/contras-enhanced(CE)-CT monotemporal images. In addition, 6 patients of head and neck cancer have a respective CE-CT multitemporal image. More specific information about this data set is presented in Appendix B.1.

### 3.1.1 Hierarchical structure selection for PET/CT

The image partition convention of $\Psi$ (max-tree, min-tree, tree of shapes, ...) must guarantee that the object of interest is defined inside the graph. The max-tree and the tree of shapes are convenient structures to consider regions with local high intensity (low intensity for a tree $\phi$) (Figure 3.1). Technically, they are anti-extensive structures that allow one to regionally process an image by not adding extra information in order to not create false contours.

In radiotherapy, both tumor active body and necrosis area (cell abnormal death) are relevant to extract information for treatment application where the max-tree and the tree of shapes are preferable.

In PET, where a tumor is generally represented by high intensities, tumor modeling as a node $N_{\mathrm{PET}_{tumor}}$ is limited by image itself restrictions: low spatial

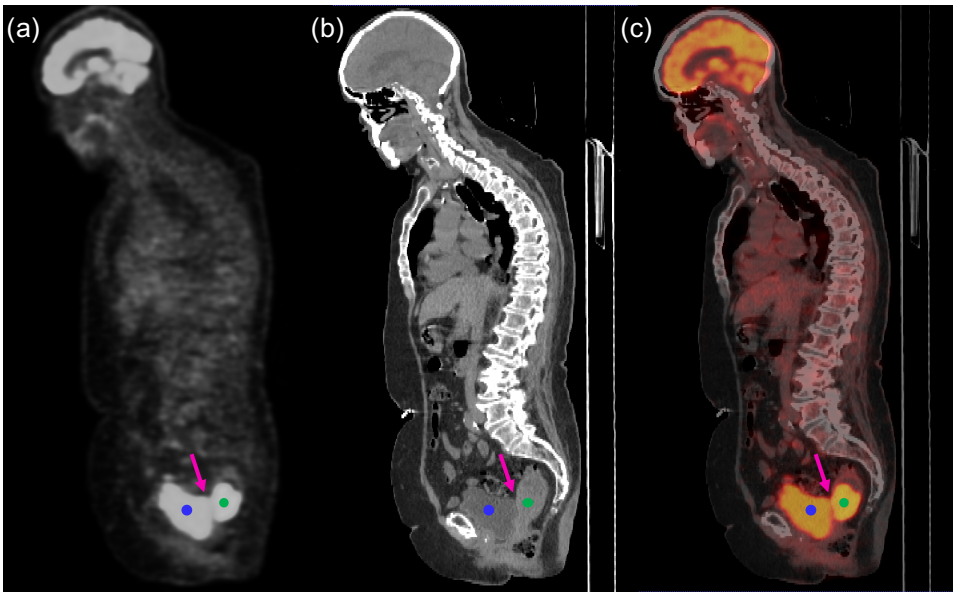


Figure 3.2: Sagittal slices of (a) PET, (b) CT and (c) fused PET/CT from a rectum cancer patient. Blue bullet is used to point to bladder zone whereas green one for rectum tumor where both show a partial overlapping (signaled by pink arrow).

resolution that leads to partial volume effect and physiological noise related to its anatomic position. A tumor can have a partial representation provoked by the adjacency with other high intensity zones (Figure 3.2). If only active tumor area is desired by the medical expert, a max-tree $\tau_{PET}^{max}$ is a better option because of ignoring low intensity intra-tumoral zones and moreover providing a lower time computing than a tree of shapes. If necrosis or cavities inside active tumor are considered for radiotheraphy, the tree of shapes is a better choice.

In CT, a tumor often has a low contrast with neighboring soft tissues, which can provoke an incomplete definition of tumor borders (Figure 3.3), which does

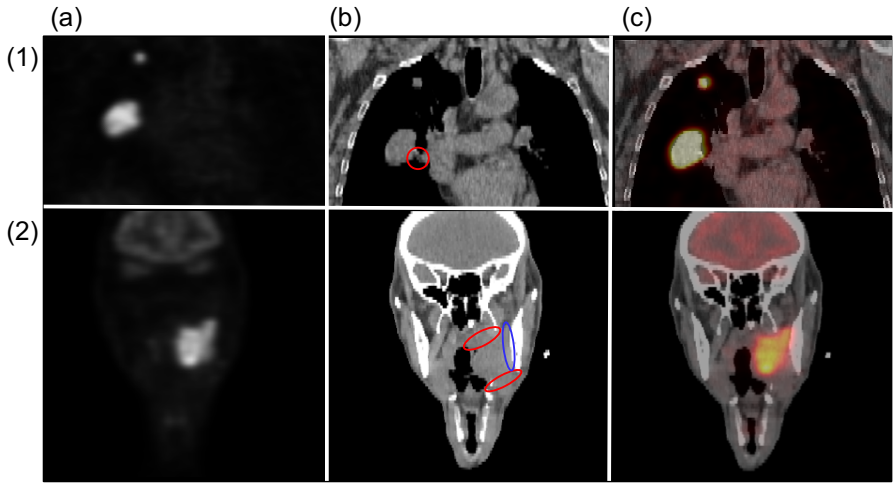


Figure 3.3: Coronal slice of (a) PET, (b) CT and (c) fused PET/CT from (1) lung and (2) head and neck cancer patients. Low contrast with soft tissues outlined by red ellipses whereas a blue ellipse or bone contact.

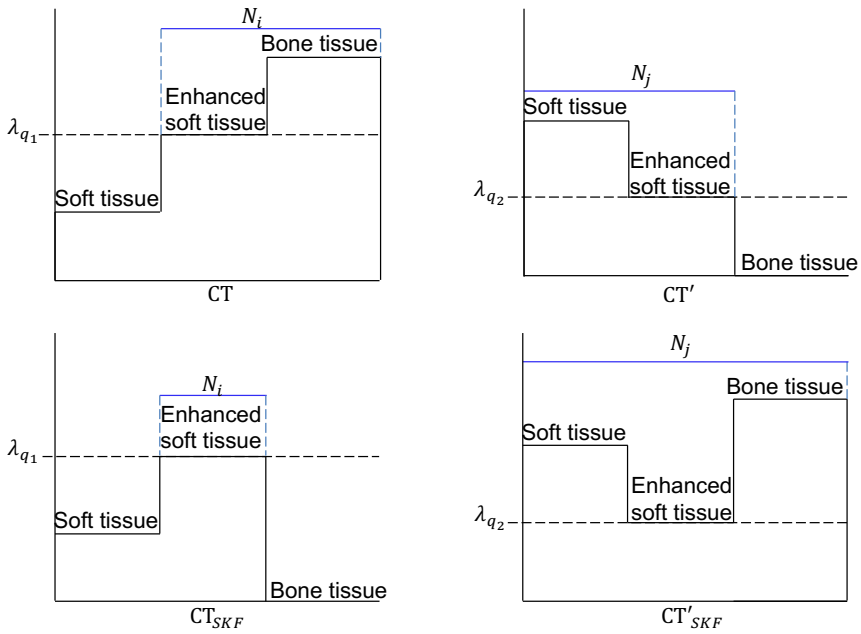


Figure 3.4: Level-set $\lambda_q$ of respective contrast-enhanced level for the partition of a node $N_i$ at CT ($\geq$ partition) and a $N_j$ at image complement CT' with and without skeleton filtering $SKF$.

not allow for its modeling as a region/node $N_{\mathrm{CT}_{tumor}}$ in the tree. A contrast agent can be intravenously injected to the patient for a contrast-enhanced CT which can improve tumor boundary definition, subject to tumor vascularization.

Moreover, soft tissues do not have the highest or lowest intensities in a CT, where bones have the highest density for a biological tissue, thus the highest intensity level (> 400 Hounsfield units (HU)); whereas air has the lowest one ($\sim -1000$ HU) [75]. Therefore, tumor node definition is limited by their potential partial contact with bones and/or air filled up cavities (Figure 3.3.(2.b)). A skeleton filtering ($SKF$) at a bone threshold $\mathrm{CT}_{SKF} = \mathrm{CT}(\Omega(q)) < 400$ HU is suitable to remove bone tissues and improve $\geq$ partition (Figure 3.4).

In general, the tree of shapes is the tree type, under presented CT inconvenients, that can have more robust representation for these possible CT versions: CT, CE-CT, $\mathrm{CT}_{SKF}$ and CE-$\mathrm{CT}_{SKF}$.

### 3.1.2 Selection of region-based attributes

Attribute selection is important for discrimination issues of tumors and their background. They are below presented according to their type and specifying the number of attributes and explaining their definition and estimation.

***Spectral attributes:*** these features are built from intensity levels $x_q$ of voxels in the node.

1. **Tree level:** characteristic intensity value $L$ coming from node level $\lambda_L$ at $\Psi$.
2. **Contrast:** it is defined as the level-set distance between a node and its parent $|N_i(\lambda_L) - N_p(N_i)(\lambda_L)|$. This metric allows to evaluate intensity changes through the branch.

***Spatial attributes:***

1. **Area**: $NoV \times Spacing3D$, where $NoV$ is the acronym of number of voxels.
2. **Barycenter coordinates**: gravity center of a node $N_i$ that is computed by the index of node voxels:

$$BC(N_i) = \frac{1}{NoV} \sum_{r=1}^{NoV} \Big[ cx_r, cy_r, cz_r \Big], \tag{3.1}$$

   where $cx, cy, cz$ are 3D coordinates of a voxel.
3. **Barycenter distance**: this parameter provides the spatial displacement of a node/region in the tree. A $dBC$ Euclidean distance between a node $N_i$ and its parent $N_p(N_i)$ barycenter is computed by:

$$dBC(N_i) = \|BC(N_i) - BC(N_p(N_i))\|. \tag{3.2}$$

***Geometric attributes***: this group of features geometrically evaluates intra-node relationships between its voxels.

1. **Boundary area:** this attribute, denoted by BA, is the $NoV$ in the boundary of a node according to an adjacency.
2. **Compactness:** also called sphericity because of comparing boundary areas (BA) of a node with that one of a synthetic sphere of same area by $\pi^{\frac{1}{3}}(6 \times \text{area})^{\frac{2}{3}}/\text{BA}$.
3. **Complexity:** parameter used to measure BA/area ratio.

***Hybrid attribute***: hybridation is applied to fuse two attribute types and build a more robust one.

1. **Spatial contrast:** this attribute, denoted SCont, weights spectral contrast according to a spatial contrast between a node and its parent by:

$$\text{SCont} = \frac{(N_p(N_i)) \setminus N_i}{N_i} \times contrast(N_i) \tag{3.3}$$

   where $contrast$ corresponds to the spectral attribute explained above.

### 3.1.3 Selection of voxel/node seeds

Voxel seeds ($x_{seeds}$) can be intuitively defined by a manual interactive selection or by automated methods. In this experimentation, we define two types of classes: foreground ($fg$), for tumor, and background ($bg$). Tumors on PET have the main characteristic of being represented by high intensity values cause of high biomarker uptake. Thus, a thresholding, denoting percentage threshold by $T$, is selected to obtain partial tumor voxels with high specificity and non-tumor zones by:

$$\begin{aligned} x_{fg} &= \text{PET}(\Omega(q)) \geq T_{fg} * \text{PET}(q_{max}), \\ x_{bg} &= \text{PET}(\Omega(q)) < T_{bg} * \text{PET}(q_{max}), \\ x_{fg} \cap x_{bg} &= \emptyset, \\ x_{seeds} &= x_{fg} + x_{bg}, \end{aligned} \tag{3.4}$$

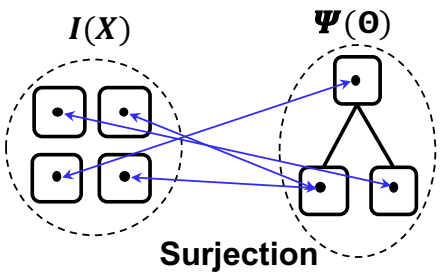


Figure 3.5: Graphic representation of surjection between voxels $\Omega$ of an image $I$ with nodes $\Theta$ of a tree $\Psi$.

where $T_{bg} < T_{fg} \in \mathbb{R}[0,1]$ to avoid $x_{seeds}$ overlapping. On the other hand, $x_{seeds}$ can be transformed by two notions related with voxel/node relationship: a node seed $N_{seed}$ and a voxel seed $\tilde{x}_{seed}$ from a tree $\Psi$.

A node seed $N_{seed}$ is a node of $\Psi$, mapped from one or more $x_{seeds}$. Concerning equation (2.8), surjective function $I(x_i) = \bigvee C_{N_j}$ is used to map a voxel $x_i$ to its respective node $N_j$ at least upper bound (lowest area, deepest subset level $\lambda_L$). Therefore, node seeds can be recovered from $x_{seeds}$ by $\bigvee C_{N_{seeds}} = I(x_{seeds})$.

Conversely, voxel seeds $\tilde{x}_{seeds}$, marked by a the symbol $\sim$ denoting tree-to-image projection $\Psi \rightarrow I$, can be obtained from $N_{seeds}$ by $I(\tilde{x}_{seeds}) = \bigvee C_{N_{seeds}}$. After these definitions, $\tilde{x}_{seeds}$ can be recovered from $x_{seeds}$ by:

$$I(x_{seeds}) \rightarrow \Psi(N_{seeds}) \rightarrow I(\tilde{x}_{seeds}), \tag{3.5}$$

where $\rightarrow$ represents $\bigvee C_N$ mapping, remarking that $x_{seeds}$ and $\tilde{x}_{seeds}$ are not necessarily equal due to voxel-to-node surjection (Figure 3.5). Indeed, depending on voxel seed selection strategy, for example, $\tilde{x}_{fg}$ from a max-tree $\tau^{max}_{\text{PET}}$ and $x_{fg}$ by equation (3.4) is a particular case where $x_{fg} = \tilde{x}_{fg}$ because of having same "equal/greater than" condition $\geq$, for $x_{fg}$ selection as well as for partition convention at $\tau^{max}$ building.

In addition, exploiting the tree structure, a hierarchical property is here added to $N_{fg}$ class, $\forall N_i$ of $\psi(N_{fg})$, $i = fg$; propagating $fg$ membership to all its descendants, potentially labeling non-marked nodes, confirming $fg$ nodes and converting $bg$ ones. Any modification at the tree can be also projected to the seeds. If a $N_{fg}$ is pruned by a filtering protocol, its voxels are replaced to the next active node parent, and thus propagating the $fg$ membership in ascendant way. In Figure 3.6, an example of $\tilde{x}_{seeds}$ propagation/repulsion is shown on a PET coronal slice of cervical cancer patient. In Figure 3.6.(c) $x_{bg}$ seeds are overlapping cavities and tumor borders due to partial low contrast as a result of heterogeneity and PVE. In Figure 3.6.(e), as a result of both filtering unstable nodes based on area region-based attribute and the tree of shapes $\phi$ configuration, $fg$ membership is ascendant and descendant sense propagated, respectively, repulsing $bg$ and covering a high part of the tumor.

### 3.1.4 Validation metric

Validation of experimental results is applied by using the Dice Similarity Coefficient (DSC) [76] as comparison metric to measure the spatial overlap (false

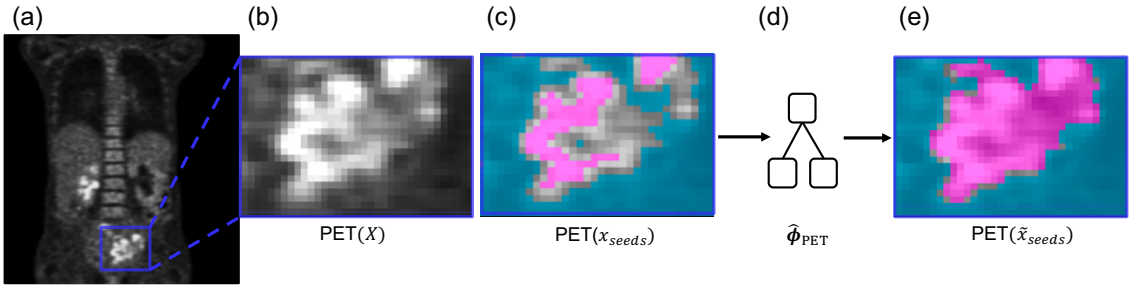


Figure 3.6: Graphic representation of seeds $x_{fg}$(pink) propagation and $x_{bg}$(blue) repulsion, to $\tilde{x}_{fg}$ and $\tilde{x}_{bg}$, through a tree $\hat{\phi}$ that was area-MSER-filtered with a $\Delta = 5$.

positives and false negatives) as follows:

$$\text{DSC}(V_{gt}, V_{seg}) = \frac{2|V_{gt} \cap V_{seg}|}{|V_{gt}| + |V_{seg}|} \tag{3.6}$$

where $V_{seg}$ is the segmented volume and $V_{gt}$ is its ground-truth. Our ground-truth is provided by an expert of Jean Godinot Institute, applying a manual segmentation on respective fused PET/CT, obtaining a binary mask of CT resolution.

## 3.2 Method 1

The method 1 (exposed in Section 2.2.1) is here applied for PET/CE-CT segmentation in a multimodal context. These medical images satisfy a spatial correspondence and high tumor representation to be evaluated in the implementation of this approach.

### 3.2.1 Experiments

In this experimentation, a random walker guided by a vectorial image is applied on 8 PET/CE-CT monotemporal data set. CE-CT is selected because of being a higher tumor-to-background contrast version and thus, coupled with skeleton filtering, increasing quality for tumor definition. Taking advantage of their spatial correspondence, from their spacing and origin position, PET is scaled to CT resolution to have a voxel-to-voxel correspondence.

PET and CE–CT$_{SKF}$ are used as input image to built potential trees: $\tau^{max}_{\text{PET}}$, $\phi_{\text{PET}}$ and $\phi_{\text{CE-CT}}$ to model respective tree combinations.

Attribute selection is equally defined for PET and CE-CT with $n = 4$ elements as: $A_{\text{PET}} = A_{\text{CE-CT}}$ = [area, BC, compacity, complexity, contrast, Scont]. This attribute selection allows to uniformly considering both data sources and the extraction of standard attributes from spatial (area and BC), geometric (compacity and complexity), spectral (contrast) and hybrid (Scont) scopes. A $\Delta = 5$ evaluation is homogeneously set for a number $n$ of $A[h]$-MSER applied to its respective $\Psi_{\text{PET}}$ and $\Psi_{\text{CE-CT}}$ to make an attribute heterogeneous analysis and obtain $n$ filtered $\hat{\Psi}_h$ at each original $\Psi_{\text{PET}}$ and/or $\Psi_{\text{CE-CT}}$. In vectorial reconstruction $I_A$, $h_*$ region-based attribute projection pertinence is evaluated. Two types of vectorial

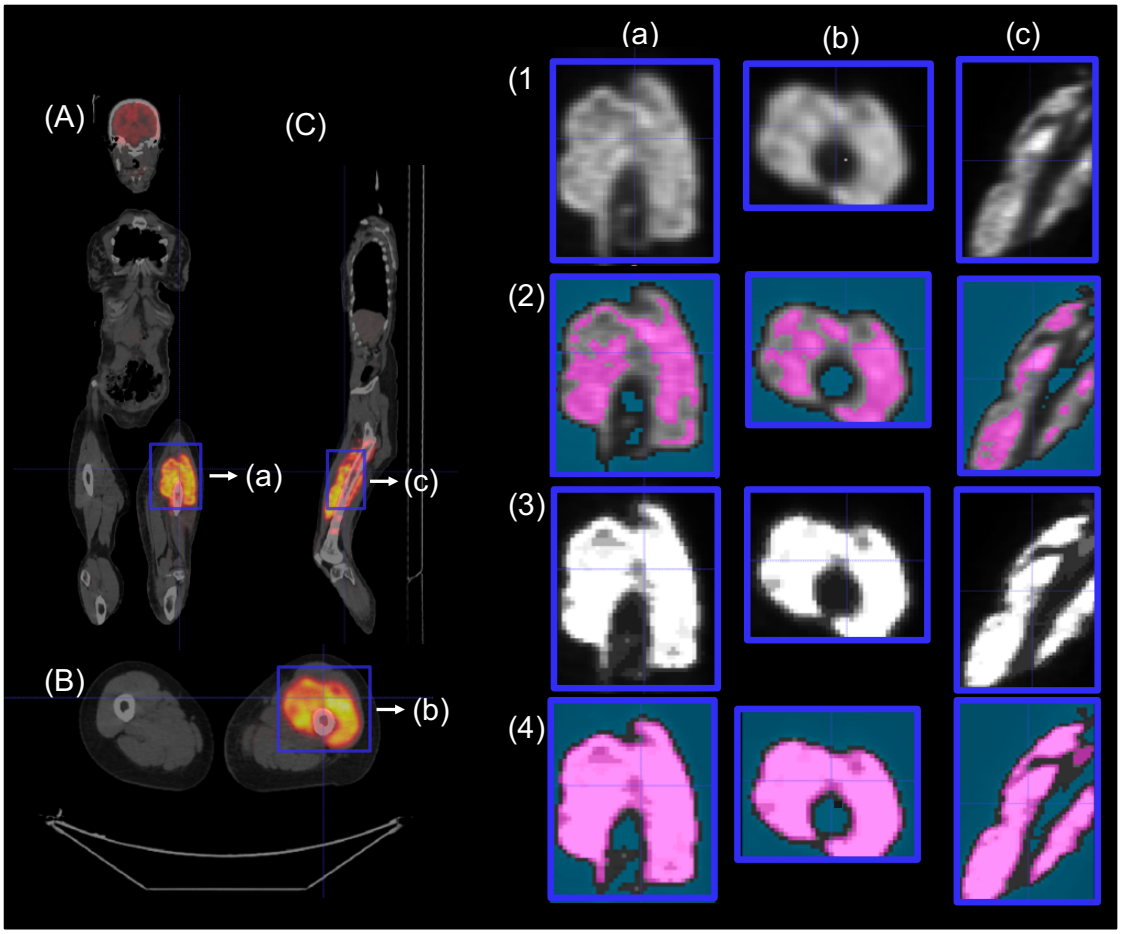


Figure 3.7: Coronal (A), axial (B) and sagittal (C) fused PET/CT (hot colormap/gray level) slices of a left thigh sarcoma patient. (a), (b) and (c) are respective cropped windows surrounding tumor. (1) PET, (2) Tumor (pink) and background (blue) seeds $v_{fg}$ and $v_{bg}$ from $T_{fg} = 0.6$ and $T_{bg} = 0.15$ percentage of $I(q_{max})$; (3) $I_{area}$ reconstructed from area-MSER filtered (1) with $\Delta = 10$ and (4) $\tilde{v}_{seeds}$ propagation from $\hat{\phi}_{PET}$.

image experiments are built by respecting projected $h_*$ from $h$-MSER: spectral and attributed vectors. If $h_* = \lambda_q$, $I_{A_q}$ is built from a heterogeneous version of $I$ as determined by $h$-MSER filtering. If $h_* = h$ that was used for $h$-MSER filtering, an attributed image $I_{A_h}$ is defined. Multimodality is represented on a vectorial image $I_A$ by concatenating, at channel of attributes, individual vectorial matrix $\text{PET}_A$ and $\text{CE-CT}_A$ for 8 channels $I_A(x) = [x_{\text{PET}}, x_{\text{CE-CT}}]$.

During the random walker step on vectorial image $I_A$, $\beta$ is set at 1 and voxel seeds $v_{seeds}$ ($x_{seeds}$ as nodes) are selected by a threshold protocol on PET because of its high tumor/background contrast (equation (3.4)), where $T_{fg}$ is set at 0.60 and $T_{bg}$ at 0.15 because of experimentally considered soft thresholds (under and upper than 0.4-0.5 classic clinical thresholds) that let our filtering and segmentation algorithms work on [0.15, 0.6] interval data (Figure 3.7).

A step of voxel seed $x_{seeds}$ propagation is here applied to project MSER filtering effect and tree configuration to voxel seeds by filtered $\hat{\Psi}$ structuring (Section 3.1.3). Using equation (3.5), by defining local parameters as $\text{PET}(x_{seeds}) \rightarrow \hat{\Psi}_{area}(N_{seeds}) \rightarrow \tilde{x}_{seeds}$, where node voxel seeds $v_{seeds} = \tilde{x}_{seeds}$ are ob-

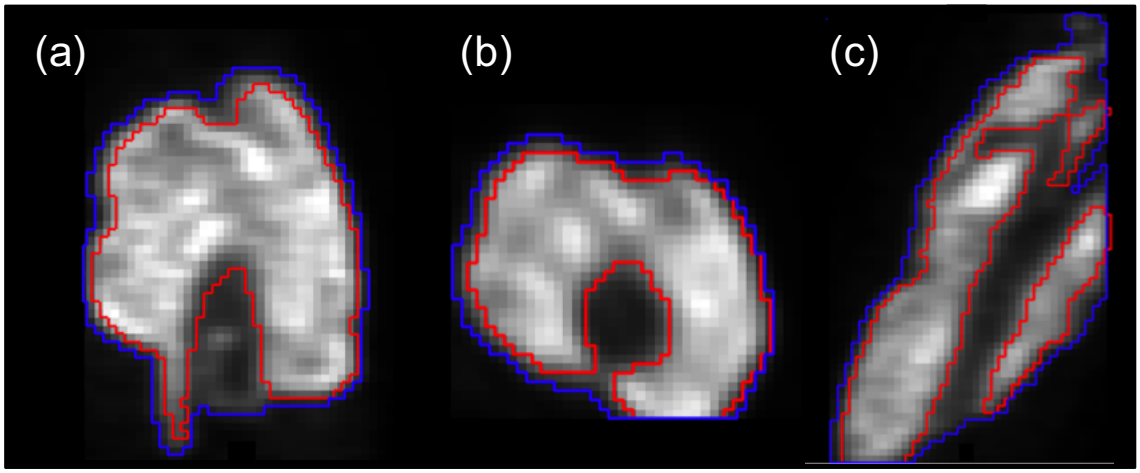


Figure 3.8: Coronal (a), axial (b) and sagittal (c) PET segmentations from a $\Delta = 5$ (red boundaries) and from a $\Delta = 15$ (blue boundaries), with a $\phi_{PET}$.

tained to start random walking. Figure 3.8 shows the impact provided by the use of trees $\phi$ for a suitable partition and the $\Delta$ coefficient as stability interval to define a holded or non-holded tumoral region.

Results (Table 3.1) were compared between vectorial and scalar random walker method applied both on monomodal images: PET and CE-CT; and multimodal PET/CE-CT image.

On scalar random walker, same $v_{seeds}$ are used as well as for vectorial random walker, differing on weighting function that is just defined by intensity level gradient (equation (2.3)). Multimodality is here interpreted, based on *Bagci et al.* [1], by fusing results $p_{\text{PET}}$ and $p_{\text{CE-CT}}$ by:

$$p_G = p_{\text{PET}}^{\alpha_1} \otimes p_{\text{CE-CT}}^{\alpha_2}, \tag{3.7}$$

where $\otimes$ refers to tensor product operation, and $\alpha_1$ and $\alpha_2$ are exponents in the

| Scalar random walker | | Vectorial random walker | | |
|---|---|---|---|---|
| $I$ combination | DSC | $\Psi$ combination | Reconstruction | DSC |
| PET | 0.66 | $\tau_{\text{PET}}^{max}$ | RW($I_{A_q}$) | 0.70 |
| | | | RW($I_{A_h}$) | 0.71 |
| CE-CT | 0.68 | $\phi_{\text{PET}}$ | RW($I_{A_q}$) | 0.73 |
| | | | RW($I_{A_h}$) | 0.73 |
| PET, CE-CT[*Bagci et al.*] | 0.69 | $\phi_{\text{CE-CT}}$ | RW($I_{A_q}$) | 0.75 |
| | | | RW($I_{A_h}$) | 0.77 |
| | | $\phi_{\text{PET}}, \phi_{\text{CE-CT}}$ | RW($I_{A_q}$) | 0.82 |
| | | | RW($I_{A_h}$) | 0.83 |

Table 3.1: Random walker (RW) results on monomodal and multimodal approaches. DSC results comparing ground-truth with vectorial RW segmentation specifying tree combinations from $\text{PET}_A$ and $\text{CE-CT}_A$ to build $I_{A_q}$ or $I_{A_h}$.

interval $[0,1] \in \mathbb{R}$ that regulate PET and CT influence in the global possibility map $p_G$. Here, both $\alpha_1$ and $\alpha_2$ are set at 0.5 to homogeneously consider PET/CT.

### 3.2.2 Conclusions

Vectorial random walker, represented by $h$-MSER filtered channels with $q$ or $h$, shows a higher performance than scalar one, based on its multi-stability attribute filtering consensus paradigm, where image regions, in PET and CT, for each attribute, has its particular evolution through the branches of a tree. Indeed, the notion of attribute value stability within hierarchical structures allows us to better discriminate relevant properties from semantic noise. The versatility of relevant attributes and imaging modalities selection, allows one to adapt the algorithm according to inherent target properties.

In addition, attributes can help to differentiate regions which, only by intensity, are considered quite similar. Results show that $A_h$ channels provide a slightly higher DSC for most of vectorial random walker experiments, arguing that attribute values provide more sensitive tumor tracks than gray intensity values $A_q$. However, the quantity of unmarked voxels is highly decreased after MSER filtering, which is shown as an important step to correct false seeds $x_{bg}$ and increase specificity of tumor seeds $x_{fg}$ by projecting the influence of filtered tree $\hat{\Psi}$ at seeds $x_{seeds}$ to $\hat{x}_{seeds}$. This seed correction impact can be measured by tumor heterogeneity where a low level produces an almost complete labeling.

Both scalar and vectorial random walker experiments show that multimodality increases accuracy performance, under condition that CT has a partial representation, which corroborates results previously obtained in the literature. Indeed, their functional/morphological complementarity allows for a better delineation of tumors. For lung tumors, this fact is more obvious due to high partial contrast between soft tissues and thorax cavity.

## 3.3 Method 2

Hierarchical graph matching filtering, described in Section 2.2.2, is here experimented on a multitemporal data set for tumor segmentation. A functional/anatomical matching paradigm is considered to find regions (tumors), that are present in both PET/CE-CT images, guided by object region-based attributes. Therefore, attribute selection is highly important for discrimination issues. The brain is chosen as an important spatial reference for a non-spatial correspondence framework. Brain detection is explained in Section 3.3.1 on both PET and CT. In Section 3.3.2, experimentation is carried out exposing respective method results.

### 3.3.1 Brain detection on PET and CT for co-spatial normalization

Brain is an organ that can be automatically detected on both imaging modalities and exploited as precious anatomical reference for non-spatially corresponding images.

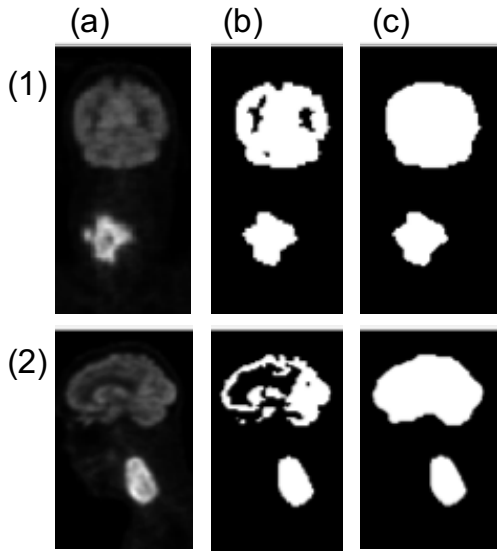


Figure 3.9: Coronal (1) and sagittal (2) PET slices (a) that are processed by (b) tresholding $T_{brain} = 0.4$ and (c) a closing with a sphere of $2cm$ radius.

The brain is expressed on FDG PET as high uptake organ, which is considered irregular because of linking to neuronal brain activity. This feature, however, can be exploited to easily segment the brain due to relatively high contrast with background. A thresholding by a percentage $T_{brain}$ of PET($q_{max}$), computed as follows:

$$\text{PET}_{bw} = \text{PET}(\Omega) > T_{brain} * \text{PET}(q_{max}), \tag{3.8}$$

can be applied on a PET to obtain a binary segmentation $\text{PET}_{bw}$ (black and white denoted $bw$).

However, uptake heterogeneity often provokes a fragmentation on the brain, including non-active cranial cavities, as ventricles, filled with cerebrospinal fluid. This problem can be solved by applying a closing operation to fuse whole brain object. A closing $\bullet$ is a morphological operation [77] known as the erosion $\ominus$ of the dilation $\oplus$ of a set $A$ by a structuring element $SE$:

$$A \bullet SE = (A \oplus SE) \ominus SE, \tag{3.9}$$

where an object $o \in O$ at $A$, such as a subset of connected voxels according to a 26-adjacency (Figure 2.1), is processed by passing a $SE$ through their borderline. Those $o$, according to a lower $SE$ size/shape spatial separation between them, are fused. Let $A = \text{PET}_{bw}$ and $SE$ be a sphere with $2cm$ radius. Between binary objects obtained by $\text{PET}_{bw} \bullet SE$, following an anatomical position, the $o_i$ placed in the highest PET at level $z$ (axial) space and/or with the greatest size, is considered as $o_{\text{PET}_brain}$ (Figure 3.9).

In Computed Tomography (CT), the brain, composed by white and gray matter (20–45 HU) interval and intracranial cavities, has a high peripheral contrast with skull but very low differentiation with spinal cord (40–50 HU) [78] which in turn is connected with several muscles (10–40 HU) and other soft tissues by peripheral nerves. A skeleton filtering $\text{CT}_{SKF} = \text{CT}(\Omega(q)) < 400$ HU is convenient to remove the skull and the spinal column as main central nervous system envelope. Remaining brain/spinal cord connection can be broken by an opening

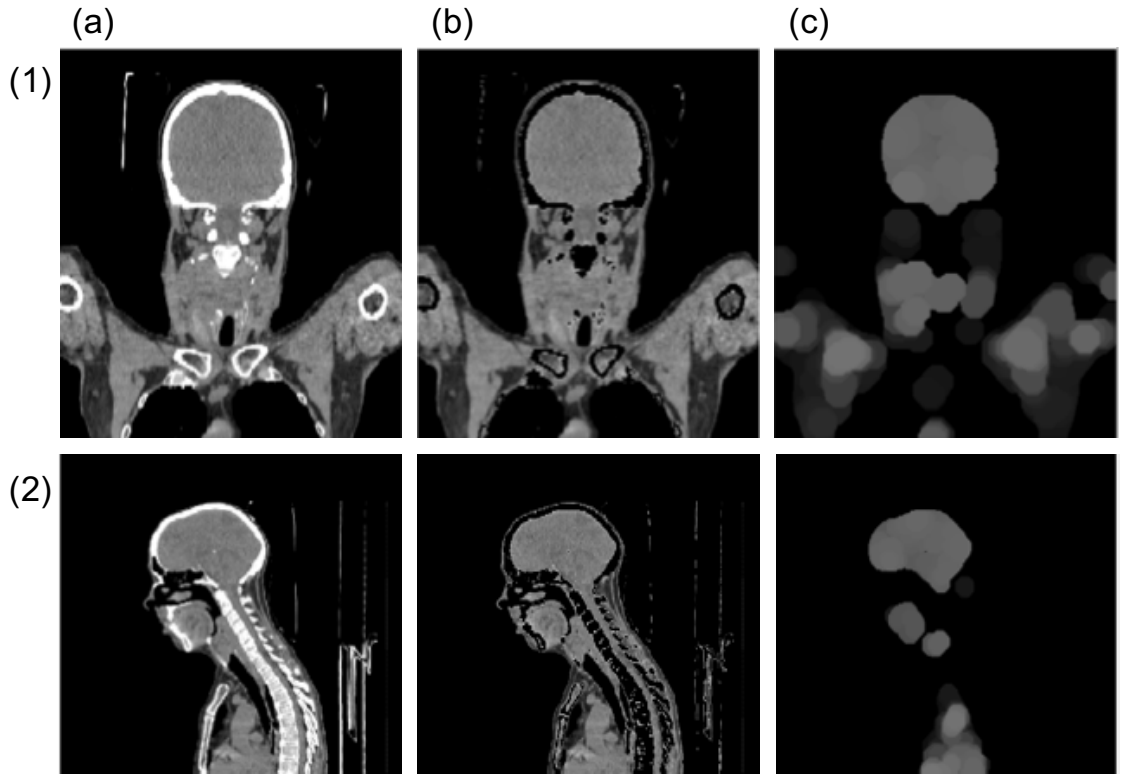


Figure 3.10: Coronal (1) and sagittal (2) CT slices (a) that are processed by (b) skeleton thresholding < 400 HU and (c) an opening with sphere of $7mm$ radius.

operation $\circ$, known as the dilation $\oplus$ of the erosion $\ominus$ of a set $A$ by a structuring element $SE$:

$$A \circ SE = (A \ominus SE) \oplus SE \tag{3.10}$$

where $A = \mathrm{CT}_{SKF}$ and $SE$ is a sphere with $7mm$ radius taking into account spinal cord $13mm$ average thick in the cervical level (Figure 3.10).

Applying a threshold on $\mathrm{CT}_{SKF} \circ SE > 20$ HU for image binarization, the $o_{\mathrm{CT}_{brain}}$ is the connected $o_i$ placed in the highest CT axial level with the greatest size.

Barycenters of both objects $o_{\mathrm{PET}_{brain}}$ and $o_{\mathrm{CT}_{brain}}$, denoted by $o_{I_{brain}}$, are used to normalize barycenter node attribute $N_i(\mathrm{BC})$ by:

$$N_i(\mathrm{BC}_{bn}) = N_i(\mathrm{BC}) - o_{I_{brain}}(\mathrm{BC}), \tag{3.11}$$

where $bn$ represents this brain normalization on barycenter attribute BC at node $N_i$.

### 3.3.2 Experiments

The multitemporal set, composed of neck and head cancer patients, is here used for experimentation. This anatomical region is characterized by high presence of bone structures as the skull, the mandible and cervical vertebrae that can perturb tumor definition. A skeleton filtering is thus applied by $\text{CE-CT}_{SKF} = \text{CE-CT}(\Omega(q)) < 400$ HU to remove bones.

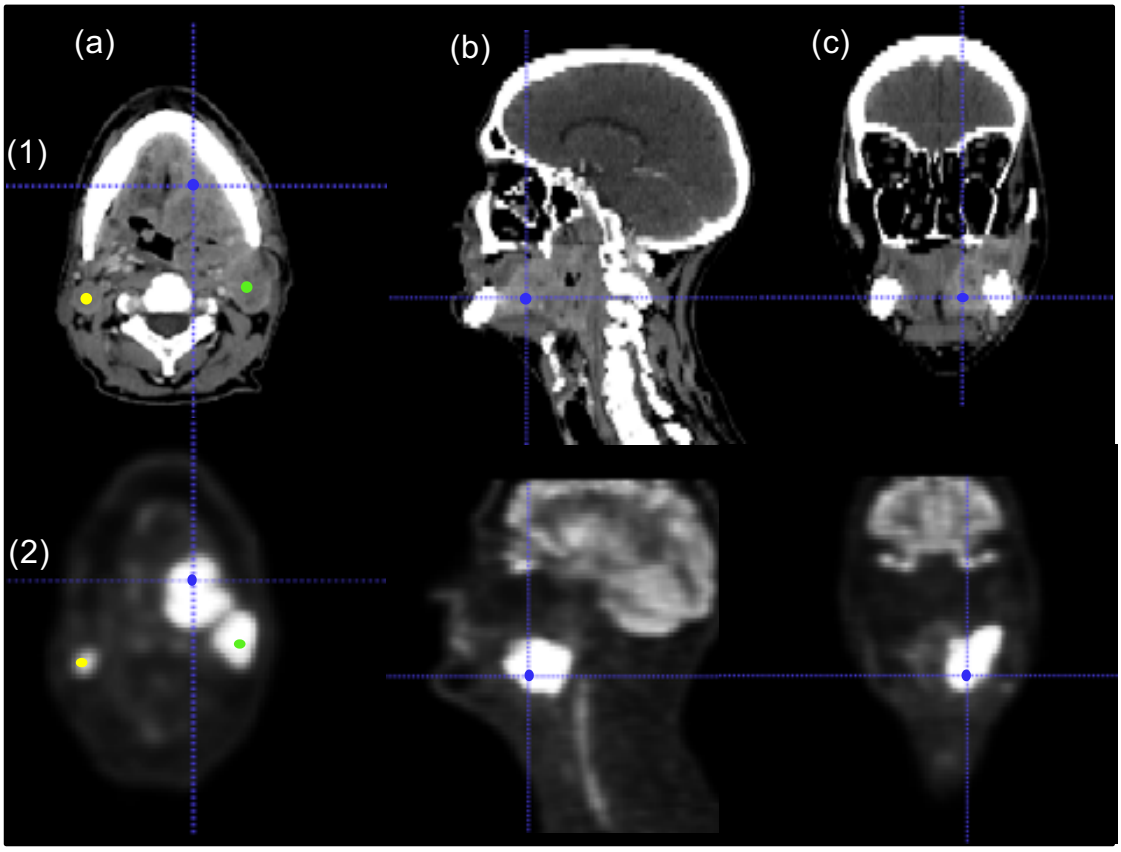


Figure 3.11: Axial (a), sagittal (b) and coronal (c) slices of CE-CT (1) and PET (2) multitemporal images of a neck and head cancer patient. Blue bullet is used to point to GTV whereas yellow and green bullets to right and left under a metastatic lesion (lymphadenopathy).

The radiotherapy department asked for a gross tumor volume (GTV) segmentation, including necrosis area, to proceed into a clinical application (Figure 3.11). A tree of shape $\phi$ is thus built from each PET and CE-CT. Selected region-based attributes must be representative of tumors for both image modalities in a global frame. Thus, $A = [\text{area, compacity, BC}_{bn}]$ comprised our selection to be tested, giving aside spectral attributes as non-comparable SUV/HU metrics. Area multiplied by respective image spacing unit (3D voxel size), makes it a comparable spatial metric; compacity as an adimensional geometric region; and finally the local brain normalized barycenter $\text{BC}_{bn}$ as geometric value.

Once the tree is built, filter processing are applied to improve the quality of data by removing noisy or non-relevant information. We take advantage of anti-extensive tree processing characteristics, not to add extra-data and just remove relatively small nodes lower than $1cm^3$ by area filtering, here considered lower than treated tumors. The weighting vector $\beta$ was set at $[0.3, 0.3, 0.4]$ giving a slightly higher importance to barycenter coordinates as spatial reference.

The lower $\varepsilon$, the lower the dissimilarity tolerance, and the number of matched nodes. This similarity threshold is set at $\varepsilon = 0.1$ in order to obtain candidates, followed of matching processing. Non-matched nodes are pruned, thus reconfiguring both filtered trees $\hat{\phi}_{\text{PET}}$ and $\hat{\phi}_{\text{CE-CT}}$. PET is used to detect nodes $N_{\text{PET}_{tumor}}$ within $\hat{\phi}_{\text{PET}}$ due to its high specificity, where $N_{\text{CE-CT}_{tumor}}$ is a co-matched node

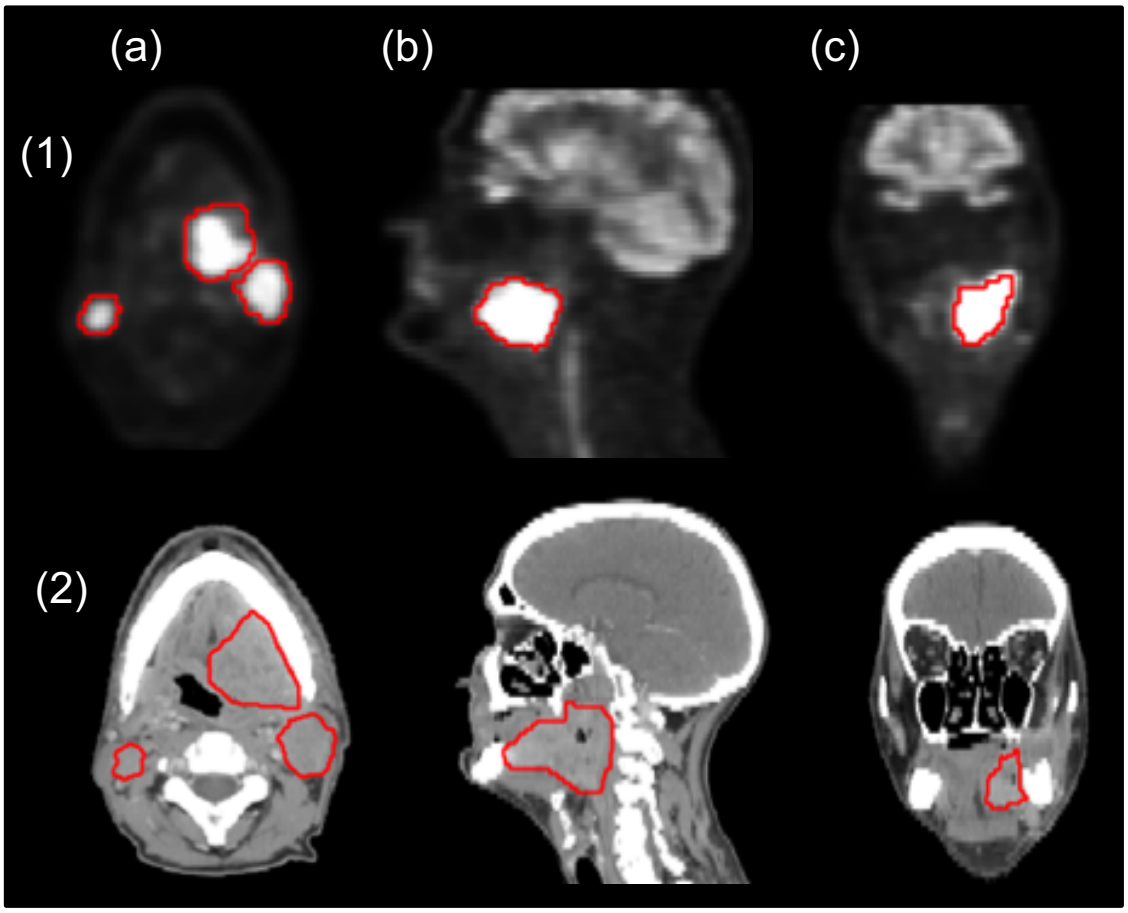


Figure 3.12: Axial (a), sagittal (b) and coronal (c) slices of PET (1) and CE-CT (2) views of segmentation result (red boundaries) in matched images.

of $N_{\text{PET}_{tumor}}$ by the matching $\Xi(N_{\text{PET}_{tumor}}, N_{\text{CE-CT}_{tumor}})$. A threshold percentage $T_{fg} = 0.6$, equation (2.12), is used to define $x_{seeds}$ that are going to be mapped to their respective active nodes $N_{PET_{fg}}$ at filtering $\hat{\phi}_{\text{PET}}$ and matched $\hat{\phi}_{\text{CE-CT}}$.

For validation, both matched tumors (Figure 3.12) were DSC evaluated with their respective CE-CT and PET ground-truth (Table 3.2). For comparison, monotemporal random walker PET versions are applied, scalar random walker with initial parameters (equivalents to those used in method 1 version): $\alpha_1$ and $\alpha_2$

| Segmentation method | Modality volume | DSC |
|---|---|---|
| Our method | PET | 0.71 |
| | CE-CT | 0.73 |
| $T_{40\%}$ | PET | 0.57 |
| RW | PET | 0.66 |
| Vectorial RW | PET | 0.81 |

Table 3.2: DSC results comparing ground-truth with our matching method, thresholding with $T_{fg} = 0.4$, scalar and vectorial random walkers (RW). Vectorial random walker presents the highest DSC, however our method allowed to project a considered relevant 0.73 DSC to a multitemporal CE-CT framework.

are set at 0.5, $T_{fg} = 0.6$ and $T_{bg} = 0.15$; while these similar initial parameters are used for vectorial random walker adding their specific method parameters: vector of attributes $A_{\text{PET}} = [\text{area, BC, compacity, complexity, contrast, SCont}]$, $\Delta = 5$ and tree propagated seeds $\tilde{v}_{seeds}$ by $\hat{\phi}_{PET}$.

### 3.3.3 Conclusions

This matching analysis is not fully topological, because child-parent relation just helps to avoid hierarchical crossing problem, preserving a hierarchy coherence, but not forcing to preserve a topological structure at node matching conditioning. Indeed, PET and CT do not have same physical origin and semantic meaning, which provides a partial co-representation of body structures, expressed on a different internal consistence. For instance, brain in PET is represented by high active part including all lobes and cerebellum whereas, for CT, the spinal cord is also connected with the brain due to low tissue contrast. Moreover, a tumor itself for a PET can have a high heterogeneity expressed by cropped gray level parts while for CT, there is not a notable difference on Hounsfield units.

Spatial correspondence handling between medical images is generally a complex issue when both information PET and CE-CT from multitemporal acquisition are precious for tumor treatment, and an image registration is considered high time-consuming and not perfect because of being prone to errors.

Tree of shapes computing seems a propitious solution for non-spatially correspondent object evaluation. Once built, this structure can be interpreted as spatially independent by node-to-node paradigm. In other words, objects of interest can be equally distributed in the graph, no matter local displacements according to a background.

Results show a greater spatial accuracy than threshold methods, but a lower one than our method 1 in PET monotemporal case applied on this section data set. However, it should be noted that tumors exhibit a high matching efficacy. Indeed, brain barycenter normalization is an important step to improve spatial analysis, taking advantage of brain co-representation and its easy detection in both PET and CT. Thus, brain barycenter distance provides an approximation to a similar spatial frame for body regions from both PET and CT, specially in those ones which are closer to the brain, as neck and head cancer this case.

## 3.4 Method 3

In this section, the method 3 (Section 2.2.3) is experimented on PET and CT images to show how to exploit these both data on a tree structure merged with the random walker in the tree as graph framework. All this, taking advantage of each image property to be used either to define regions than to take part as contextual information. These two possible roles have their interest to be selected based on the their impact that can be expressed in the results.

### 3.4.1 Experiments

Monomodal and multimodal scopes are analyzed in local method experiments in our whole data set of 33 medical images in monotemporal. PET, CT and CE-

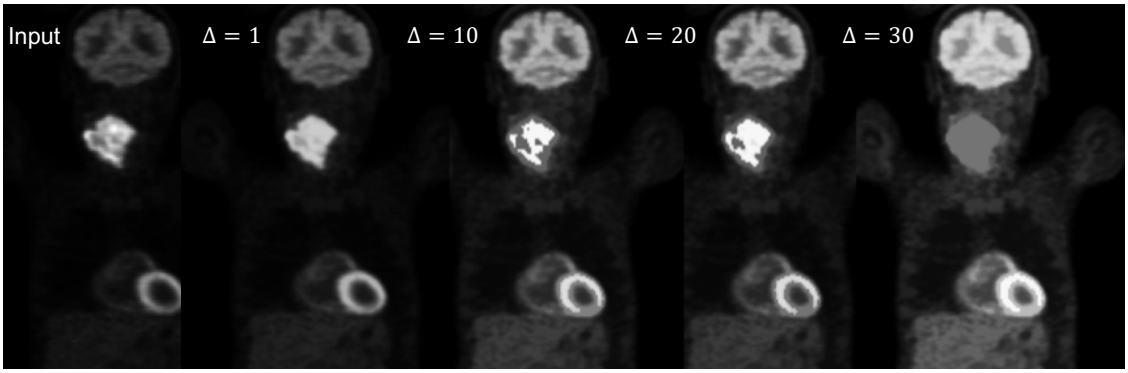


Figure 3.13: Majority vote (MV)-MSER $\phi(\text{PET}, -)$ evolution with $\Delta[1, 10, 20, 30]$. For visualization, respective $\lambda$ intensity value is assigned in image reconstruction process, having a percentage of $[98.32\%, 99.17\%, 99.53\%, 99.67\%]$ of pruned nodes.

CT are the potential inputs to be used at two possible roles: $I_{\text{base}}$ for tree building and $I_{\text{context}}$ as complementary information at the tree, denoted by $\Psi(I_{\text{base}}, I_{\text{context}})$. The symbol – is applied to represent not $I_{\text{context}}$ usage. From $I_{\text{context}}$, a new type of region based attributes is presented: context information attributes (Section 2.1.2).

Region-based attributes are selected according to image role at the tree, building a vector $A = [A_{\text{base}}, A_{\text{context}}]$. From $I_{\text{base}}$ data, $A_{\text{base}} = [\text{area, BC, contrast, compacity}]$ are chosen because of standard spatial, spectral and geometric attributes; whereas from $I_{\text{context}}$, $A_{\text{context}} = [\text{contrast}_{\text{context}}]$ to measure contextual information variation.

A majority vote MSER filtering is then applied on $\Psi$ with a $\Delta$ set at 1 as a standard parameter base on a $\Delta$ evaluation. In Figure 3.13, $\Delta$ node distance impact is shown on majority vote MSER filtered $\phi(\text{PET}, -)$. In average, a 98.32% of nodes are pruned after majority vote MSER with the lowest $\Delta = 1$, showing an efficiency to delete non-relevant data, considerably reducing processing data and

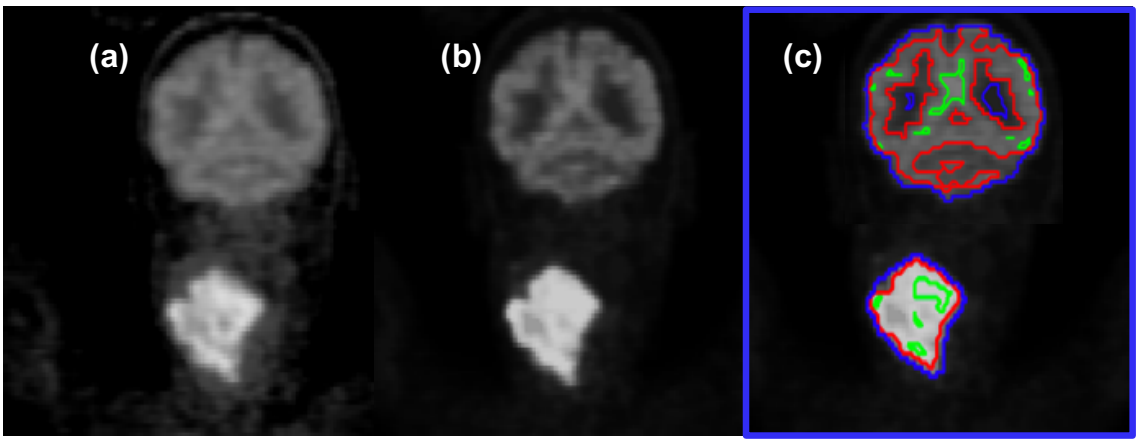


Figure 3.14: PET coronal slice of head and neck cancer patient (a). Majority vote (MV)-MSER PET with $\Delta = 1$ (b). (c) Green boundaries represents voxel seeds labeled by $T_{fg}$, red and blue boundaries represent $N_{fg}$ and $N_{bg}$ node seeds mapped from their respective voxel seeds.

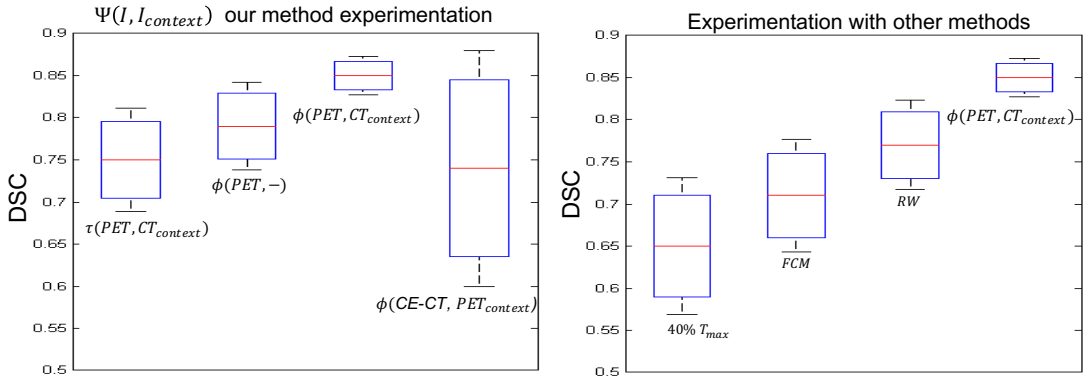


Figure 3.15: DSC results obtained in PET/CT and PET/CE-CT clinical experimentation of our methodology with different combinations, and other methods of state of the art such as 40%$T$ threshold on $q_{max}$, fuzzy-c-means (FCM) and PET/CT random walker (RW).

preserving relevant regions such as tumors and high uptake organs. A decreasing intensity is obvious since once pruned, regions get a lower $\lambda$ level in the tree.

Two classes are here defined, foreground (tumor) as positive class $c^{+}_{fg}$ and background as passive class $c^{-}_{bg}$.

A test varying $T_{fg}$ and $T_{bg}$ between [0.10, 0.30] and [0.50, 0.75], respectively, was done to analyze area attribute sensitivity between mapped pixel/voxel $x_{fg}$ and $x_{bg}$ to corresponding $N_{fg}$ and $N_{bg}$ on majority vote MSER filtered $\phi$(PET, $-$). $N_{fg}$(area) variates ±9% and $N_{bg}$ with ±17% showing a reduction on seed definition dependency. This behavior is sustained by MSER pruning that suppressed noisy data, decreased heterogeneity and provoked a tree graph rearrangement including intensity distribution in a decreasing order. For random walker segmentation on $\Psi$, $T_{fg}$ and $T_{bg}$ are set in 0.15 and 0.6 (Figure 3.14), while $\beta[h]$ is homogeneously fixed on $1/H$ not giving relevance for a special feature.

In the obtained results (Figure 3.15), $\phi$(PET,CT) got the highest average DSC, arguing that CT improves accuracy sustained in anatomical references. On the other hand, $\phi$(CE-CT,PET) highlight in 3 images of HNC test getting a higher accuracy of 0.88 DSC but a great standard deviation and a lower average DSC. Indeed, this partial accuracy is highly dependent on tumor vascularization for contrast flow, which is often irregular. If a tumor volume is anatomically well represented on a CE-CT, PET data projection on its tree of shapes model, asserts a high performance for identification and segmentation.

Our data set was also experimented with state of the art methods: a classic threshold of 40% on $I(q_{max})$; a fuzzy-c-means (tumor and background classes); and PET/CT random walker with $x_{fg}$ and $x_{bg}$ as seeds $v_{seeds}$; where our method on $\phi$(PET,CT) remains the most accurate.

### 3.4.2 Conclusions

This methodology allows to take advantage of our whole data set of 33 images without restrictions. The two first methods require the use of CT or CE-CT if

the respective tumor was well defined under the partition type of selected tree (max-tree or tree of shapes). This requirement is no longer required in the third method where CT is used as contextual information, allowing to take benefits but having non-negative consequences on the result. CT has a low contrast tumor to background which is often not enough to build a regional representation by a tree of shapes. However, CT data taken as a contextual information modeled on a PET tree, can be an important feature to guide tumor segmentation. Indeed, soft tissues are difficult to be identified between them; however, their partial contact with bones and air filled up cavities can provide precious information to detect partial borders.

Experiments with contrast-enhanced CT as tree based showed a high standard deviation due to physiological issues related with irregular contrast product flow inside tumors. However, a shiny high accuracy is shown obtaining the most partial accurate results.

PET modeled as a tree of shapes and injected with CT contextual information was the most accurate average experimentation. These results support the fact that multimodality can improve segmentation performance. This fact is coherent since ground-truth is delineated on a fused bimodal scope and not only based on PET.

Our methodology provides a strong integration between their main components to make it robust: a tree structure as framework; and majority vote MSER and random walker, as filtering and segmentation for image processing. The max-tree and the tree of shapes are convenient structures to treat high intensity image regions as tumors. Technically, they are anti-extensive structures that allow to regionally treat an image by not adding extra information. Majority vote MSER shows their advantages to smooth out fuzzy borders, not to treat low relevant nodes such as noise and thus getting a lighter structure. This allows to get a faster (since 98% lower time cost) and optimal segmentation process. The random walker, as segmentation methodology, is only applied on uncertainty branch sections (not all graph) by taking advantage of tree structural qualities. This methodological coupling, based on arguments mentioned above and results obtained on clinic experimentation, seems to be robust face to PET/CT tumor segmentation.

# Conclusions et perspectives

Les trois méthodologies exposées dans ce travail partagent l'implémentation d'une structure hiérarchique $\Psi$ afin de représenter convenablement les régions spectrales d'une image. Premièrement, ceci peut être justifié en raison de l'interprétation spectrale d'une tumeur, principalement en TEP, sous forme de hauts niveaux de gris locaux $q$ positionnés vers les feuilles de l'arbre. Deuxièmement, ceci permet d'analyser l'évolution des nœuds à travers les branches jusqu'à la racine. Finalement, ceci permet d'exploiter les informations locales d'un nœud. En effet, la caractérisation des nœuds avec des attributs basés sur les régions peut favoriser le suivi des nœuds d'intérêt tels que les tumeurs pour la segmentation ou les nœuds sans intérêt associés au bruit exprimé en des petites régions à enlever.

La multimodalité TEP / TDM est ici exploitée pour la segmentation, soit pour tirer directement parti des deux images en tant que structures hiérarchiques, soit pour utiliser indirectement leur contenu dans le processus de segmentation. En effet, l'idée de modéliser une image en tant qu'arbre est fondée sur le prémisse que l'objet d'intérêt peut être représenté sous la forme d'un nœud.

La TEP fournit un contraste tumeur-fond plus élevé pour détecter et pour représenter une tumeur en tant que nœud $N$ alors que la tomodensitométrie présente un faible contraste face aux autres tissus mous. On obtient néanmoins à partir de ces sources d'information une définition complète de la tumeur. Cependant, leur contact partiel avec les cavités ou les os peut fournir des informations pertinentes pouvant être utilisées indirectement dans un arbre, comme indiqué dans la méthode 3. De plus, l'injection d'un produit de contraste iodé pour la IC-TDM peut améliorer le contraste tumoral selon la pathologie et la vascularisation. L'amélioration du signal peut alors améliorer la définition de la tumeur au niveau du nœud, ce qui est le cas de la méthode 2, ainsi qu'un voxel lui-même par rapport à son voisinage dans l'espace de la matrice, ce qui est exploité dans les méthodes 1 et 3.

L'arbre de formes $\phi$ a été retenu comme étant une structure de données fournissant une configuration performante pour représenter des objets tels que des tumeurs avec des cavités ou une nécrose. Sur les trois méthodes, cette structure hiérarchique a obtenu les performances les plus élevées par rapport au max-tree $\tau$, au prix d'un temps de calcul plus long. Cependant, ces méthodes diffèrent par la temporalité des images d'entrée et les composants de l'algorithme.

Les méthodes 1 et 3 se concentrent toutes les deux sur la TEP / TDM monotemporelle en tirant parti de leur correspondance spatiale. En revanche, la méthode 2 repose sur un cas multitemporel, où la TEP et la IC-TDM sont acquises à différentes dates. Le but est alors de trouver une solution à la non-correspondance spatiale en exploitant l'arbre à partir de sa structuration hiérarchique et son or-

ganisation structurée mais non-spatialisée.

La méthode 1 et la méthode 3 sont étroitement liées en raison de la composition similaire avec une arborescence structurant l'image et un marcheur aléatoire utilisé lors de la segmentation. Toutefois, dans la méthode 1, l'arborescence est simplement utilisée comme complément, tandis que dans la méthode 3, elle constitue la structure principale du processus de segmentation. Dans la méthode 1, un arbre est utilisé pour le filtrage MSER afin d'augmenter le contraste entre les nœuds et d'éliminer les données non stables (y compris les données bruitées), puis de projeter les valeurs d'intensité / d'attributs dans l'espace image pour une segmentation par marche aléatoire vectorielle. Dans la méthode 3, une version discriminante de MSER par vote majoritaire est appliquée lors du filtrage, suivie d'une adaptation du marcheur aléatoire lors de la segmentation complète de l'arbre évaluant l'évolution de la tumeur jusqu'à la racine.

# Perspectives

Les travaux futurs pourront se concentrer sur les points suivants :

**–Augmenter le niveau de validation**

- augmenter la taille de la base de données TEP / TDM dans les portées multitemporelles (du même patient) et multimodales (TEP / TDM, TEP / IRM)
- ; analyser de manière approfondie chaque attribut régional pour déterminer son impact individuel ;
- étudier la pertinence des attributs régionaux (nouvelles combinaisons) afin d'augmenter le niveau de discrimination pour la caractérisation de la tumeur en tant qu'objet d'intérêt ;
- appliquer ces méthodes à d'autres domaines de recherche tels que la dosimétrie d'organes ;
- utiliser d'autres mesures de validation telles que l'indice de Jaccard afin d'évaluer simultanément la similarité et la dissimilarité des résultats de l'algorithme par rapport à la vérité terrain.

**–Améliorer les qualités intrinsèques de nos méthodes**

- augmenter la robustesse de la mise en correspondance des arbres par la sélection aléatoire des candidats dans la méthode 2.

**–Travaux potentiels**

Ce travail est dédié à la précision de la segmentation. D'autres projets peuvent être générés pour tirer parti de ces résultats de segmentation :

- classifier les zones de forte intensité afin de discriminer automatiquement les zones où la fixation est anormale de celles où elle est normale ; ainsi que identifier les zones de transit des biomarqueurs dans les vaisseaux sanguins et le tractus urinaire ;

D'autres pistes méthodologiques peuvent être explorées pour améliorer ou lancer des projets de recherche :

- fusion des principaux avantages de la méthode 2 (multitemporalité) et de la méthode 3 (précision) afin de produire une méthode entièrement multimodale et multitemporelle (performances supérieures) ;
- algorithme d'appariement topologique à partir de deux arbres de deux images multitemporelles avec une cohérence topologique (tels que deux TDM multitemporelles possédant une cohérence anatomique).

# Conclusions and perspectives

The three methodologies exposed on this work share the implementation of a hierarchical structure $\Psi$ in order to conveniently process the spectral regions of an image. Firstly, this can be justify because of spectral interpretation of a tumor, mainly in PET, as local high gray levels $q$ that are placed towards the leaves of the tree. Secondly, this allows to analyze node evolution through branches way to the root. Finally, this allows to exploit local node information. Indeed, node characterization with region-based attributes can promote the tracking of nodes of interest such as tumors for its segmentation, or non-interest nodes as noise corresponding to small regions to be removed.

PET/CT multimodality is here exploited for segmentation, either to directly take advantage of both images as hierarchical structures, as well as indirectly using their information in the segmentation process. Indeed, the idea of modeling an image as a tree is sustained under the premise that the object of interest can be represented as a node. PET provides a higher contrast tumor-to-background to detect and represent a tumor as a node $N$; whereas CT has a low contrast with soft tissues. However, their partial contact with cavities and bones can provide rich information that can be indirectly used in a tree, as shown in Method 3. Moreover, iodine contrast injection for CE-CT can improve tumor contrast according to respective pathology and vascularization. This enhancement quality can improve tumor definition either as node, which is a mandatory condition for Method 2, as well as a voxel itself with its neighborhood at image matrix space, available for Method 1 and Method 3.

The tree of shapes $\phi$ has been shown as a relevant data structure to represent min/max objects such as tumors with cavities or necrosis. For the three methods, this hierarchical structure got the highest performance compared with the max-tree $\tau$ under the constraint of a higher time computing. However, they differ in input image and algorithmic components. Indeed, Method 1 and Method 3 both focus on monotemporal PET/CT, taking advantage of their spatial correspondence, whereas Method 2 focuses on non-spatial correspondence case, as PET and CE-CT acquired on different studies, by exploiting tree graph linking, partition and non-spatial essence over node-to-node paradigm. Method 1 and Method 3 are highly related because of similar composition with a tree as image structure and random walker segmentation. However, in Method 1, the tree is only used as a complementary source of information, whereas in Method 3, it is the main structure in the segmentation process. In Method 1, a tree is used for MSER filtering to increase the contrast between nodes and remove non-stable data and then project intensity/attribute values to matrix space for a vectorial random walk segmentation. For Method 3, a higher discriminant version of MSER by majority vote is applied for filtering, followed by a random walker full-tree segmentation

evaluating tumor evolution way to the root.

## Perspectives

In future works, the following points could be explored:

**–To enhance validation level**

- a larger PET/CT database in multitemporal (from same patient case) and multimodal scopes (PET/CT, PET/MRI);
- a deeper analysis of region-based attributes to determine their individual impact;
- a research about the relevance of region-based attributes (new combinations) in order to increase discrimination level for tumor characterization;
- methodological application in other research projects as organ dosimetry;
- other validation metrics as Jaccard index to evaluate either similarity than dissimilarity of algorithm results face to a ground-truth.

**–To improve intrinsic method qualities**

- robustness increasing of tree matching (method 2) by random candidate selection.

**–Potential works**

This work is focused on segmentation accuracy. Other projects can be generated to take advantage of these results:

- a classification of high intensity zones is planned to be studied to find patterns and automatically discriminate abnormal from normal high uptake zones and those belonging to biomarker transit as blood vessels and urinary tract.

Other methodological paths can be followed to improve or start research projects:

- fusion of main advantages of method 2 (multitemporality) and method 3 (accuracy) to produce a fully multimodal and multitemporal method (higher performance);
- topological matching algorithm from two trees of two multitemporal images with a topological coherence (as two multitemporal CT by their anatomic coherence).

# Appendix A

# Publications and communications related with the thesis work

My scientific production, throughout my doctoral formation, can be summarized on:

- Francisco J. Alvarez-Padilla, Barbara Romaniuk, Benoît Naegel, Stéphanie Servagi-Vernat, Dimitri Papathanassiou and Nicolas Passat. *Matching filtering by region-based attributes on hierarchical structures for image co-segmentation.* The 25th IEEE International Conference on Image Processing (ICIP), Megaron Athens International Conference Centre, Athens, Greece, October 7–10, 2018, pp. 131–135, DOI: 10.1109/ICIP.2018.8451384.

- Francisco J. Alvarez-Padilla, Barbara Romaniuk, Benoît Naegel, Stéphanie Servagi-Vernat, David Morland, Dimitri Papathanassiou and Nicolas Passat. *Hierachical forest attributes for multimodal tumor segmentation on FDG-PET/contrast-enhanced CT.* The 15th IEEE International Symposium on Biomedical Imaging (ISBI), Omni Shoreham Hotel, Washington D.C., United States of America, April 4–7, 2018, pp. 163–167, DOI: 10.1109/ISBI.2018.8363546.

- Francisco J. Alvarez-Padilla, Barbara Romaniuk, Benoît Naegel, Stéphanie Servagi-Vernat, Dimitri Papathanassiou and Nicolas Passat. *Les arbres de coupes comme extracteur d'attributs pour la segmentation multi-images par marche aléatoire.* The 26th GRETSI colloquium, Palais des Congrès, Juan-Les-Pins, France, September 5–8, 2017.

- Francisco J. Alvarez-Padilla, B. Romaniuk, B. Naegel, S. Servagi-Vernat, D. Papathanassiou and N. Passat. *Segmentation multi-images et multicritères par arbres de coupes*. Journée des Jeunes Chercheurs, SFR CAP-Santé, Reims, France, Mars 30, 2017.

A preliminary work related with this thesis, on my MSc internship, produced the paper below:

- Francisco J. Alvarez-Padilla, Éloïse Grossiord, Barbara Romaniuk, Benoît Naegel, Camille Kurtz, Hugues Talbot, Laurent Najman, Romain Guillemot, Dimitri Papathanassiou and Nicolas Passat. *Multicriteria 3D PET image segmentation*. The 5th International Conference on Image Processing Theory,

Tools & Applications (IPTA), Orléans, France, November 10–13, 2015, pp. 346–351, DOI: 10.1109/IPTA.2015.7367162.

# Appendix B

# Medical image specificities

## B.1 Medical data-set

A total of 33 patient image sets were recovered from Jean Godinot Institute (Reims, France) which are distributed as follows:

- 12 lung cancer,
- 8 head and neck cancer,
- 3 breast cancer,
- 3 lymphoma,
- 2 hepatic cancer,
- 3 cervix cancer,
- 1 colon cancer,
- 1 sarcoma.

The 8 head and neck cancer PET/CE-CT images are implemented for Methods 1 and 2, whereas the whole data set is implemented for Method 3. A set of 8 multitemporal CE-CT images, corresponding to PET/CE-CT head and neck cancer patients, is used for Method 2 experimentation.

### 25 PET/CT and 8 PET/CE-CT

At nuclear medicine department, acquisitions were carried out by using a Discovery 710 PET/CT scanner (General Electrics). Patients received an intravenous dose of $^{18}F$-FDG (3 MBq/kg). Portal phase CT was performed first in auto mA mode with adaptive statistical iterative reconstruction, native collimation of 16×1.25 mm, resolution matrix of $512 \times 512 \times 343$ with a voxel size of $0.97 \times 0.97 \times 2.5$ mm$^3$. PET acquisition was performed 60 minutes after FDG injection, covering the area from the base of the skull through the proximal femur. PET was reconstructed using an iterative algorithm (OSEM 24 subsets, 2 iterations) and post-filtered with a Butterworth filter (cut-off frequency: 6.4 mm), obtaining a $256 \times 256 \times 263$ matrix with a voxel size of $2.73 \times 2.73 \times 3.27$ mm$^3$.

For the case of $^{18}F$-FDG PET/CE-CT, an intravenous dose of iodinated contrast agent (Optiject 350mg/mL) was injected 70 seconds before CT phase.

## 8 CE-CT

Radiotheraphy departement provided 6 multitemporal CE-CT corresponding to 6 oral cancer patients with a respective PET/CT monotemporal set. CE-CT were carried out by using a Brilliance CT Big Bore scanner (Philips). Intravenous dose of iode solution (IOMERON 300mg/mL) was injected in two fractions, 70% and 30% having two time-out laps after their injection, 60 and 30 seconds, to proceed to CT acquisition. It was performed in an helical mode, with a collimation of 16×0.75 mm, 120 kV, 350 mAs with a sharp filter. Reconstruction was carried out by a filtered backprojection providing a resolution matrix of $512 \times 512 \times 258$ with a voxel size of $1.98 \times 1.98 \times 3$ mm$^3$, where axial space covers the area from diaphragm to the top of the skull.

## B.2 Medical image overview on selected patient tumors

Some PET/CT images from experimental data set are chosen to be shown and explained (Figures B.1, B.2, B.3, B.4, B.5), where (1) axial, (2) coronal and (3) sagittal slices of (a) PET, (b) CT and (c) fused PET/CT of different cancer patients, are displayed. Pointed boundary boxes encase high intensity regions related with (red) abnormal high uptake, (yellow) normal high uptake and (blue) biomarker transit zones. White arrows for PET and green ones for CT are used to point out respiratory motion zones, respectively. In the caption of each image, their specific information is detailed.

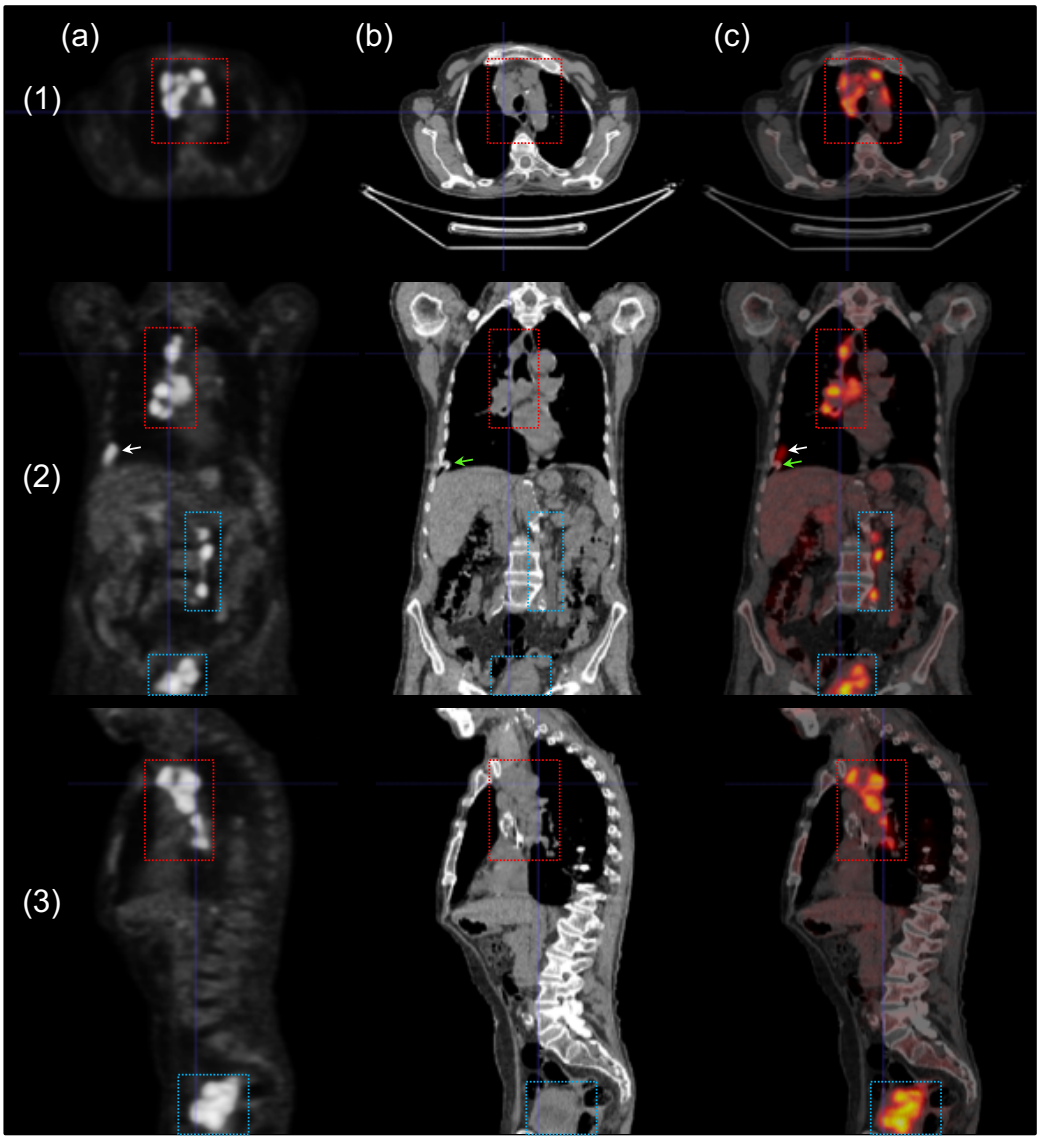


Figure B.1: Lung cancer patient with a cancerous nodule in the right inferior lobe and multiple mediastinal adenopathies (2). Right inferior lobe shows in (2.c) a volume spreading at (2.a), greater tumor volume than (2.b), due to respiratory motion, which is really common in diaphragm proximity region. In addition, this local non-spatial correspondence can result in an underestimation of SUV from an incorrect attenuation correction based on CT.

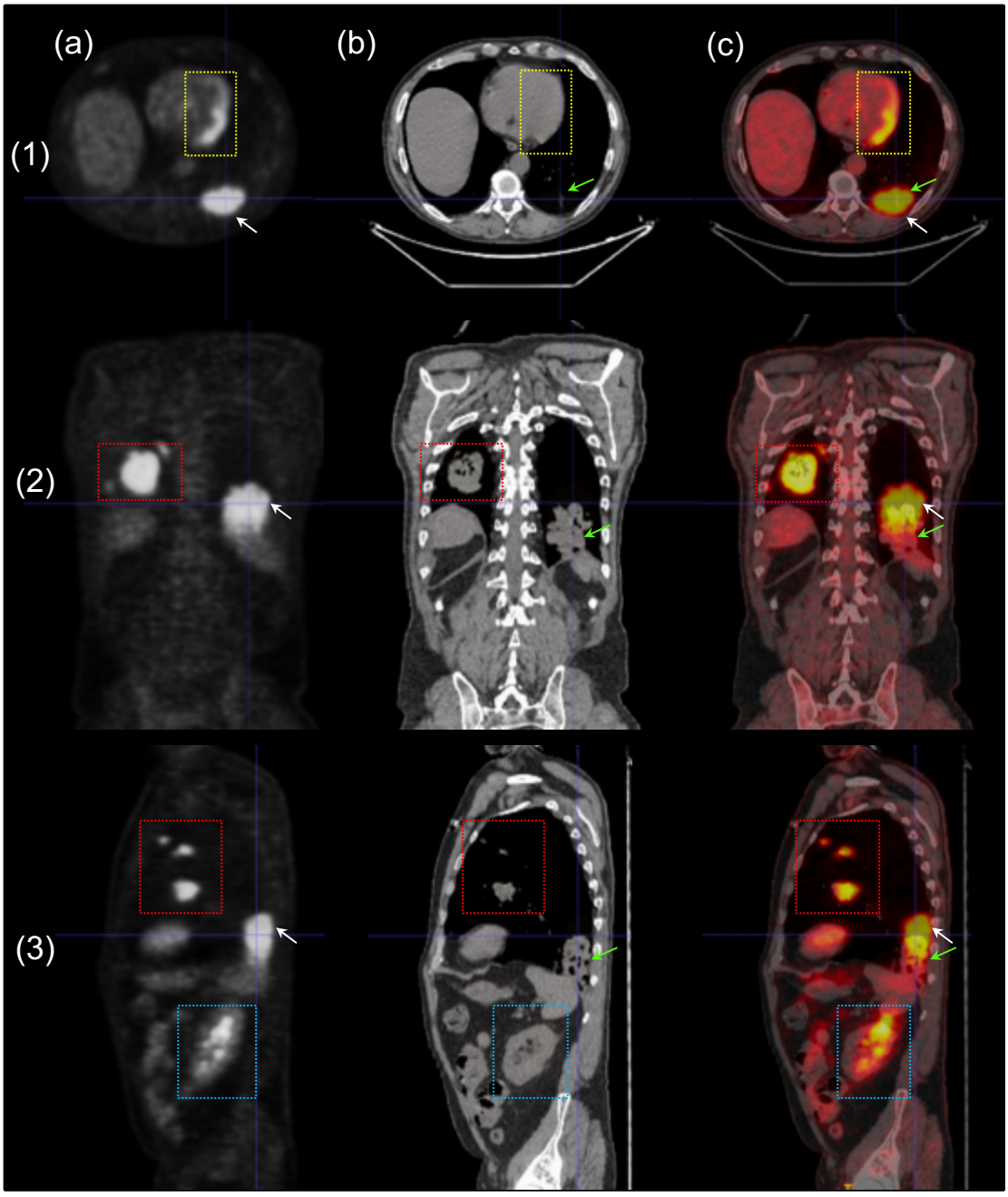


Figure B.2: Bilateral lesions in a lung cancer patient (1,2,3). Left lesion (2), which is close to the diaphragm, shows a tumor spreading, which is obvious in (1.b) where lesion is not present for CT whereas for (1.a) PET, it is. In (3), we observe a high intensity at left kidney because of biomarker transit.

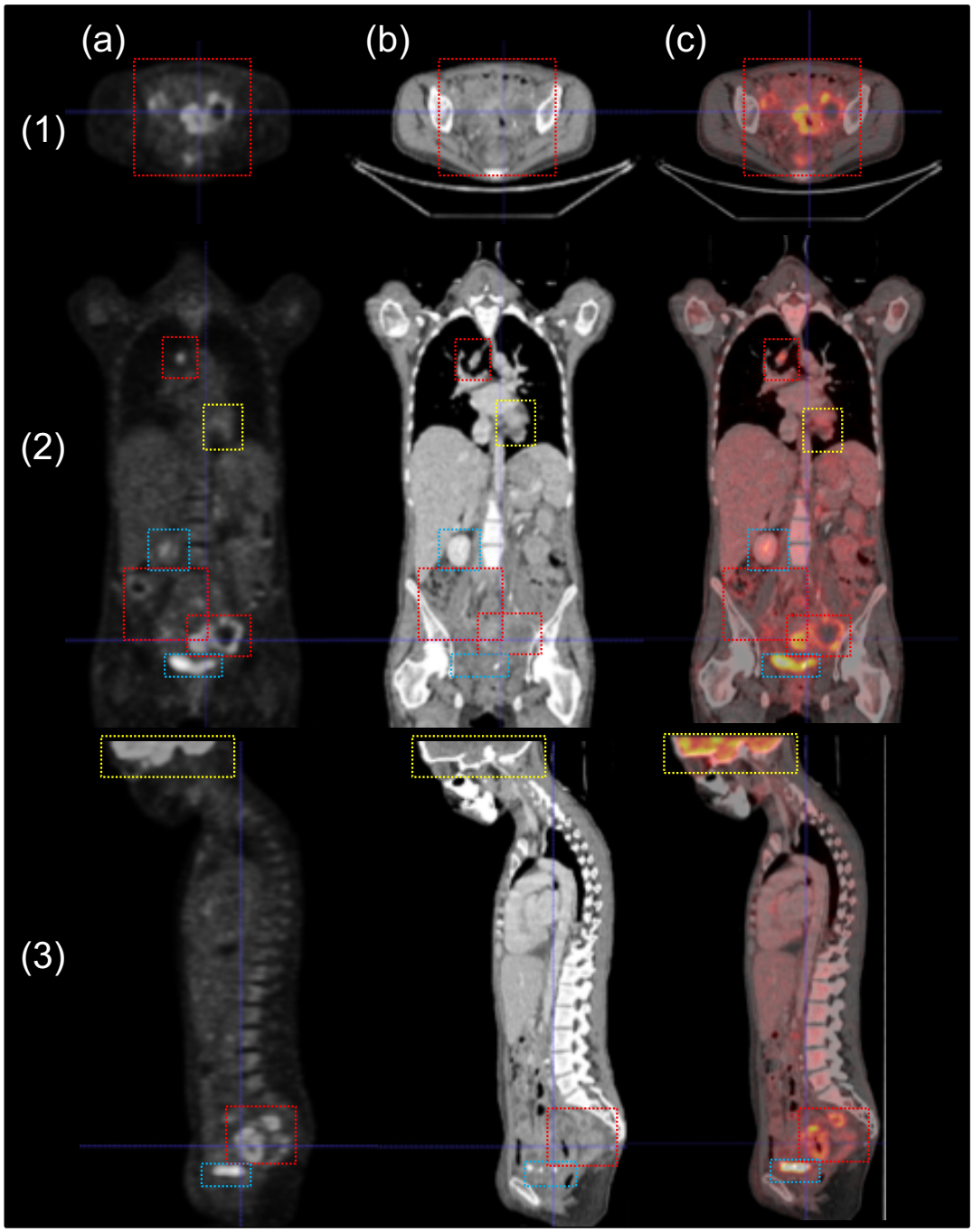


Figure B.3: Patient with cervical cancer of the uterus with peritoneal and thoracic ganglia metastasis. Brain and myocardium normal high uptake are shown in (2) and (3). A highly disseminated peritoneal carcinomatosis is observed in the right side of (2). In (2,3), the bladder and right kidney show radioactivity transit.

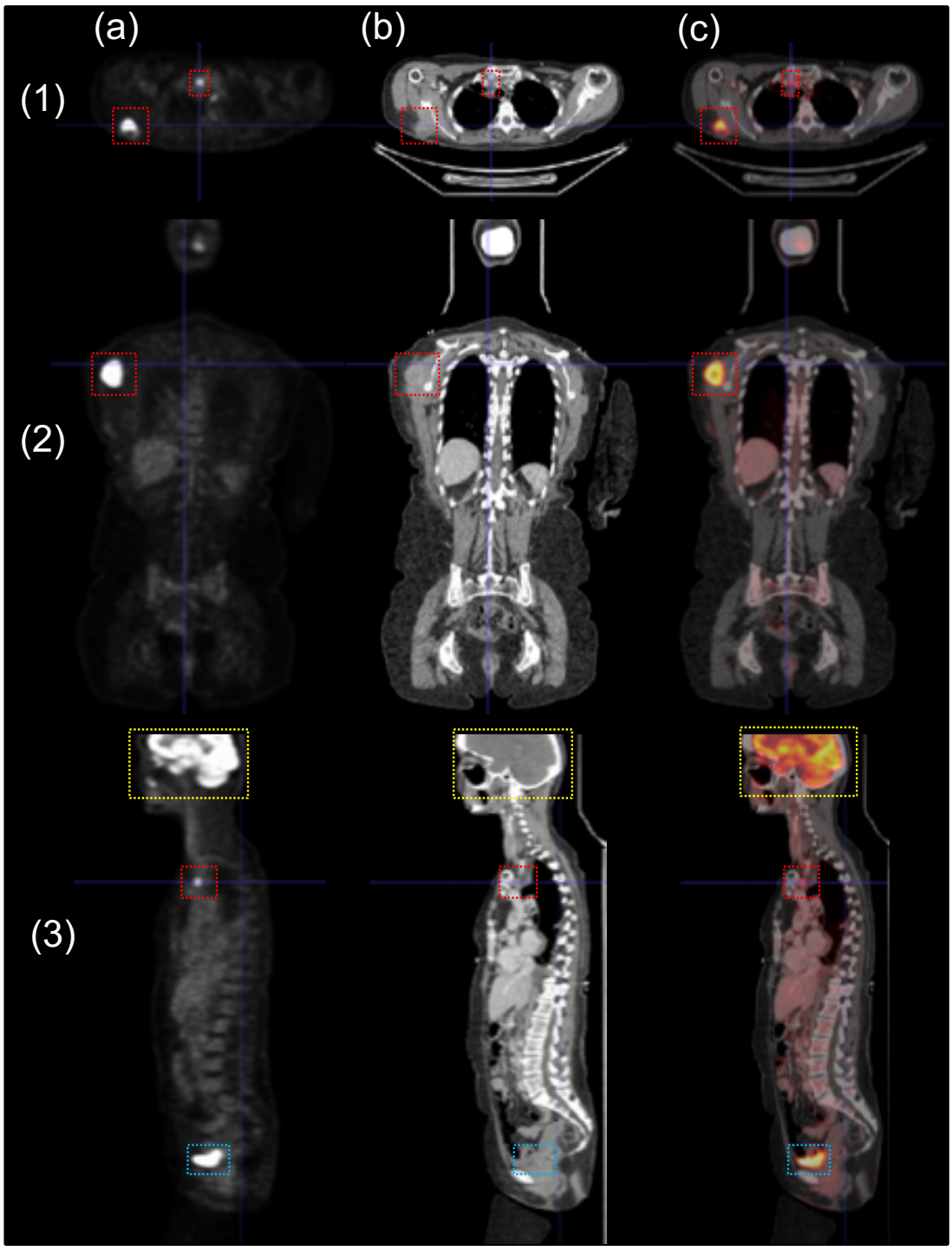


Figure B.4: Supradiaphragmatic lymphoma with dorsal (posterior to the scapula) (1,2) and mediastinal involvement (1,3). Brain (3) with a normal high uptake from high activity neurons and bladder as last radioactive urine deposit (3).

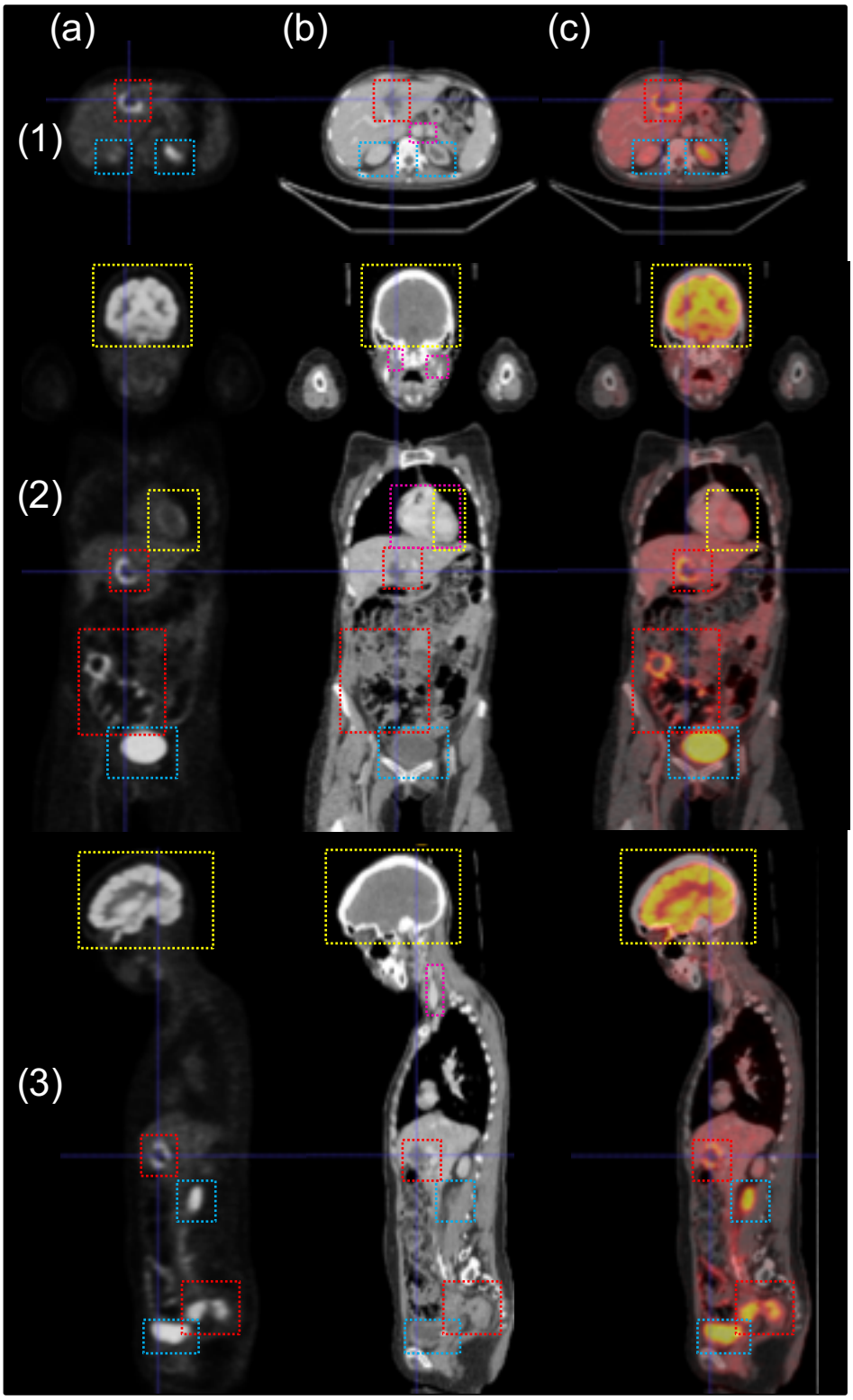


Figure B.5: Uterus cervix cancer with hepatic and peritoneal metastasis. Cancerous liver segment IV is shown in (1, 2, 3). In the right side of (2), we observe peritoneal lesions, and in (3), upper than the bladder, lesions in the pelvis. Brain (2,3) and the myocardium (2) show a normal high uptake. Both kidneys in (1), right kidney in (3) and the bladder in (2,3), show the presence of radiotracer transit. In (b), contrast enhanced CT can be identified by high blooded zones where contrast agent is flowing. Looking at pink dashed boxes, there is a slightly higher intensity in the inferior aorta and renal branches in (1.b), as well as in the heart as blood pump (2.b) and neck vessels in (2.b) and (3.b).

## B.3 DICOM-RT structure production for IMRT

With regard to the project with the radiotherapy department at Jean Godinot Institute, leaded by Stéphanie Servagi-Vernat, results obtained from neck and head tumor segmentation experiments in Section 3.3, were evaluated in order to plan intensity modulated radiation therapy (IMRT) [79]. This type of radiotheraphy treatment seeks to adapt the radiation dose according to tumor heterogenous activity, by increasing the dose intensity under aggressiveness molecular level, generally called dose painting [80]. To this end, binary mask of segmented tumor is evaluated based on SUV from original PET intensities by a sub-fragmentation according to $n$ subintervals at [PET($q_{max}$)–PET($q_{min}$)].

From each sub-region, voxels that are non-fully connected according to the 26 adjacency, are considered as boundary voxels. These borders, on matrix space, are translated into DICOM-RT Structure [81] as a standard format for practical management in clinical field, Figure B.6. These DICOM-RT files were sent to radiotherapy department for validation and then to compute modulated dose at each sub-tumor zone and thus proceed to IMRT clinical application.

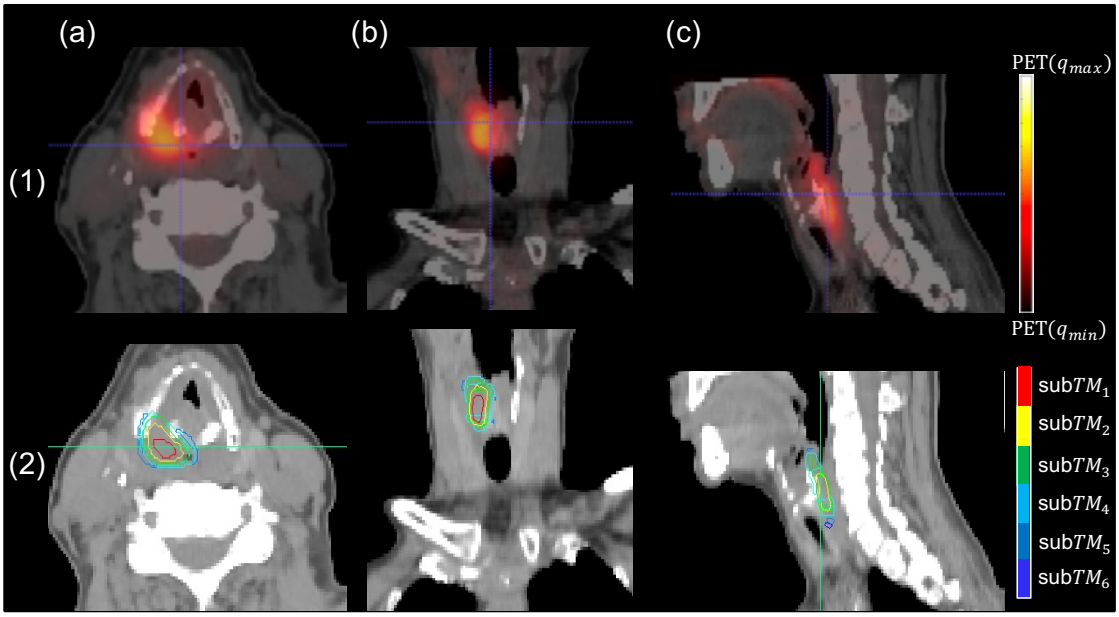


Figure B.6: (a) Axial, (b) coronal and (c) sagittal slices of (1) fused PET-(hot colormap)/CE-CT and (2) CE-CT with sub-tumor boundaries of 6 fragmented intervals from minimum to maximum tumor intensity.

# Bibliography


[1] U. Bagci, J. Udupa, N. Mendhiratta, B. Foster, Z. Xu, J. Yao, X. Chen, and D. Mollura, "Joint segmentation of anatomical and functional images: Applications in quantification of lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT images," *Medical Image Analysis*, vol. 17, no. 8, pp. 929–945, 2013.

[2] Y. Song, J. Bai, D. Han, S. Bhatia, W. Sun, W. Rockey, J. Bayouth, J. Buatti, and X. Wu, "Optimal co-segmentation of tumor in PET-CT images with context information," *IEEE Transactions on Medical Imaging*, vol. 32, no. 9, pp. 1685–1697, 2013.

[3] F. Bray, J. Ferlay, I. Soerjomataram, R. Siegel, L. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[4] J. Ferlay, M. Colombet, I. Soerjomataram, T. Dyba, G. Randi, M. Bettio, A. Gavin, O. Visser, and F. Bray, "Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018," *European Journal of Cancer*, vol. 103, pp. 356–387, 2018.

[5] G. Gatta, L. Botta, and M. Sánchez, "Prognoses and improvement for head and neck cancers diagnosed in Europe in early 2000s: The EUROCARE-5 population-based study," *European Journal of Cancer*, vol. 51, no. 15, pp. 2130–2143, 2015.

[6] R. Siegel, K. Miller, and A. Jemal, "Cancer statistics, 2017," *CA: A Cancer Journal for Clinicians*, vol. 67, no. 1, pp. 7–30, 2017.

[7] W. Röntgen, "Ueber eine neue art von strahlen," *Annalen der physik*, vol. 300, no. 1, pp. 12–17, 1898.

[8] G. Hounsfield, "Computerized transverse axial scanning (tomography). Part I: Description of system. Part II: Clinical applications," *British Journal of Radiology*, vol. 46, pp. 1016–1022, 1973.

[9] D. Harwood-Nash, "Computed tomography of ancient Egyptian mummies," *Journal of Computer Assisted Tomography*, vol. 3, no. 6, pp. 768–773, 1979.

[10] D. Nedden, R. Knapp, K. Wicke, W. Judmaier, W. Murphy, H. Seidler, and W. Platzer, "Skull of a 5,300 year old mummy: Reproduction and investigation with CT guided stereolithography," *Radiology*, vol. 193, pp. 269–272, 1994.

[11] W. Murphy, D. zur Nedden, P. Gostner, R. Knapp, W. Recheis, and H. Seidler, "The iceman: Discovery and imaging," *Radiology*, vol. 226, no. 3, pp. 614–629, 2003.

[12] F. Wrenn, M. Good, and P. Handler, "The use of positron-emitting radioisotopes for the localization of brain tumors," *Science*, vol. 113, no. 2940, pp. 525–527, 1951.

[13] G. Brownell and C. Burnham, "A multi-crystal positron camera," *IEEE Transactions on Nuclear Science*, vol. 19, no. 3, pp. 201–205, 1972.

[14] W. M. Townsend, D. W., L. Byars, A. Geissbuhler, H. Tochon-Danguy, A. Christin, M. Defrise, D. Bailey, S. Grootoonk, and A. Donath, "A rotating PET scanner using BGO block detectors: Design, performance and applications," *The Journal of Nuclear Medicine*, vol. 34, no. 8, pp. 1367–1376, 1993.

[15] T. Beyer, D. Townsend, T. Brun, P. Kinahan, M. Charron, R. Roddy, J. Jerin, J. Young, L. Byars, and R. Nutt, "A combined PET/CT scanner for clinical oncology," *The Journal of Nuclear Medicine*, vol. 41, no. 8, pp. 1369–1379, 2000.

[16] P. Kinahan, D. Townsend, T. Beyer, and D. Sashin, "Attenuation correction for a combined 3D PET/CT scanner," *Medical Physics*, vol. 25, no. 10, pp. 2046–2053, 1998.

[17] W. Weber, S. Ziegler, R. Thodtmann, A.-R. Hanauske, and M. Schwaiger, "Reproducibility of metabolic measurements in malignant tumors using FDG PET," *The Journal of Nuclear Medicine*, vol. 40, no. 11, pp. 1771–1777, 1999.

[18] S. Rosenbaum, T. Lind, and G. Antoch, "False-positive FDG PET uptake—The role of PET/CT," *European Radiology*, vol. 16, no. 5, pp. 1054–1065, 2006.

[19] R. Boellaard, "Standards for PET image acquisition and quantitative data analysis," *Journal of Nuclear Medicine*, vol. 50, no. 5, pp. 11–20, 2009.

[20] J. Ollinger, "Model-based scatter correction for fully 3D PET," *Physics in Medicine and Biology*, vol. 41, no. 1, pp. 153–176, 1996.

[21] S. Vandenberghe, E. Mikhaylova, E. D'Hoe, P. Mollet, and J. Karp, "Recent developments in time-of-flight PET," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 3, no. 3, pp. 1–30, 2003.

[22] M. Nemallapudi, S. Gundacker, P. Lecoq, E. Auffray, A. Ferri., A. Gola, and C. Piemonte, "Sub-100 ps coincidence time resolution for positron emission tomography with LSO:Ce codoped with Ca," *Physics in Medicine & Biology*, vol. 60, no. 12, pp. 4635–4649, 2015.

[23] P. Kinahan and J. Rogers, "Analytic 3D image reconstruction using all detected events," *IEEE Transactions on Nuclear Science*, vol. 36, no. 1, pp. 964–968, 1989.

[24] H. Hudson and R. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp. 601–609, 1994.

[25] M. Osman, C. Cohade, Y. Nakamoto, and R. Wahl, "Respiratory motion artifacts on PET emission images obtained using CT attenuation correction on PET-CT," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 30, no. 4, pp. 603–606, 2003.

[26] G. Goerres, E. Kamel, T.-N. Heidelberg, M. Schwitter, C. Burger, and G. Schulthess, "PET-CT image co-registration in the thorax: Influence of respiration," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 29, no. 3, pp. 351–360, 2002.

[27] K. Zasadny and R. Wahl, "Standardized uptake values of normal tissues at PET with 2-[fluorine-18]-fluoro-2-deoxy-d-glucose: Variations with body weight and a method for correction," *Radiology*, vol. 189, no. 3, pp. 847–850, 1993.

[28] V. Ouellet, A. Routhier-Labadie, W. Bellemare, L. Lakhal-Chaieb, E. Turcotte, C. A., and D. Richard, "Outdoor temperature, age, sex, body mass index, and diabetic status determine the prevalence, mass, and glucose-uptake activity of 18F-FDG-detected BAT in humans," *The Journal of Clinical Endocrinology & Metabolism*, vol. 96, no. 1, pp. 192–199, 2011.

[29] H. Young, R. Baum, U. Cremerius, K. Herholz, O. Hoekstra, A. Lammertsma, J. Pruim, and P. Price, "Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: Review and 1999 EORTC recommendations," *European Journal of Cancer*, vol. 35, no. 13, pp. 1773–1782, 1999.

[30] W.-D. Heiss, K. Wienhard, R. Wagner, H. Lanfermann, A. Thiel, K. Herholz, and U. Pietrzyk, "F-Dopa as an amino acid tracer to detect brain tumors," *European Journal of Cancer*, vol. 37, no. 7, pp. 1180–1182, 1996.

[31] B. Foster, U. Bagci, A. Mansoor, Z. Xu, and D. Mollura, "A review on segmentation of positron emission tomography images," *Computers in Biology and Medicine*, vol. 50, pp. 76–96, 1994.

[32] W. Yu, X. Fu, Y. Zhang, J. Xiang, L. Shen, G. Jiang, and J. Chang, "GTV spatial conformity between different delineation methods by 18 FDG PET/CT and pathology in esophageal cancer," *Radiotherapy and Oncology*, vol. 93, no. 3, pp. 441–446, 2009.

[33] D. Schinagl, W. Vogel, A. Hoffmann, J. van Dalen, W. Oyen, and J. Kaanders, "Comparison of five segmentation tools for 18f-fluoro-deoxy-glucose–positron emission tomography-based target volume definition in head and neck cancer," *International Journal of Radiation Oncology Biology Physics*, vol. 69, no. 4, pp. 1282–1289, 2007.

[34] D. Hellwig, T. Graeter, D. Ukena, A. Groeschel, G. Sybrecht, H.-J. Schaefers, and C.-M. Kirsch, "18F-FDG PET for mediastinal staging of lung cancer: Which SUV threshold makes sense?," *Journal of Nuclear Medicine*, vol. 48, no. 11, pp. 1761–1766, 2007.

[35] H. Vees, S. Senthamizhchelvan, R. Miralbell, D. Weber, O. Ratib, and H. Zaidi, "Assessment of various strategies for 18F-FET PET-guided delineation of target volumes in high-grade glioma patients," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 36, no. 2, pp. 182–193, 2009.

[36] A. Schaefer-Schuler, S. Kremp, D. Hellwig, C. Rübe, C. Kirsch, and U. Nestle, "A contrast-oriented algorithm for FDG-PET-based delineation of tumour volumes for the radiotherapy of lung cancer: Derivation from phantom measurements and validation in patient data," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 35, no. 11, pp. 1989–1999, 2008.

[37] U. Nestle, A. Schaefer-Schuler, S. Kremp, A. Groeschel, D. Hellwig, C. Rübe, and C.-M. Kirsch, "Target volume definition for 18F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 34, no. 4, pp. 453–462, 2007.

[38] Y. Erdi, O. Mawlawi, S. Larson, M. Imbriaco, H. Yeung, R. Finn, and J. Humm, "Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding," *Cancer*, vol. 80, no. 12, pp. 2505–2509, 1997.

[39] W. Jentzen, L. Freudenberg, E.-G. Einsig, M. Heinze, W. Brandau, and A. Bockish, "Segmentation of PET volumes by iterative image thresholding," *The Journal of Nuclear Medicine*, vol. 48, no. 1, pp. 108–114, 2007.

[40] J. Daisne, M. Sibomana, A. Bol, T. Doumont, M. Lonneux, and V. Grégoire, "Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: Influence of reconstruction algorithms," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 69, no. 3, pp. 247–250, 2003.

[41] Q. Black, I. Grills, L. Kestin, C. Wong, J. Wong, A. Martinez, and D. Yan, "Defining a radiotherapy target with positron emission tomography," *International Journal of Radiation Oncology Biology Physics*, vol. 60, no. 4, pp. 1272–1282, 2004.

[42] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[43] V. Caselles, F. Catté, T. Coll, and F. Dibo, "A geometric model for active contours in image processing," *Numerische Mathematik*, vol. 66, no. 1, pp. 1–31, 1993.

[44] J. Sethian, *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials science*, vol. 3. Cambridge University Press, 1999.

[45] H. Li, W. Thorstad, K. Biehl, R. Laforest, Y. Su, K. Shoghi, E. Donnelly, D. Low, and W. Lu, "A novel PET tumor delineation method based on adaptive region-growing and dual-front active contours," *Medical Physics*, vol. 35, no. 8, pp. 3711–3721, 2008.

[46] X. Geets, J. Lee, A. Bol, M. Lonneux, and V. Grégoire, "A gradient-based method for segmenting FDG-PET images: Methodology and validation," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 34, no. 9, pp. 1427–1438, 2007.

[47] M. Wanet, J. Lee, B. Weynand, M. De Bast, A. Poncelet, V. Lacroix, E. Coche, V. Grégoire, and X. Geets, "Gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer: A comparison with threshold-based approaches, CT and surgical specimens," *Radiotherapy and Oncology*, vol. 98, no. 1, pp. 117–125, 2011.

[48] M. Hatt, C. Cheze le Rest, A. Turzo, C. Roux, and D. Visvikis, "A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET," *IEEE Transactions on Medical Imaging*, vol. 28, no. 6, pp. 881–893, 2009.

[49] M. Hatt, C. Cheze le Rest, P. Descourt, A. Dekker, D. De Ruysscher, M. Oellers, P. Lambin, O. Pradier, and D. Visvikis, " Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications," *International Journal of Radiation Oncology, Biology, Physics*, vol. 77, no. 1, pp. 301–308, 2010.

[50] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, 1967.

[51] J. Tou and R. Gonzalez, "Pattern recognition principles," *Journal of Applied Mathematics and Mechanics*, vol. 57, no. 6, pp. 353–354, 1977.

[52] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *Journal of Cybernetics*, vol. 3, no. 3, pp. 32–57, 1973.

[53] J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press, 1981.

[54] A. Dewalle-Vignion, N. Betrouni, N. Makni, D. Huglo, J. Rousseau, and M. Vermandel, "A new method based on both fuzzy set and possibility theories for tumor volume segmentation on PET images," *IEEE International Conference on Engineering in Medicine and Biology Society*, pp. 3122–3125, 2008.

[55] M. Vermandel, N. Betrouni, C. Taschner, C. Vasseur, and J. Rousseau, "From MIP image to MRA segmentation using fuzzy set theory," *Computerized Medical Imaging and Graphics*, vol. 31, no. 3, pp. 128–140, 2007.

[56] S. Vial, D. Gibon, C. Vasseur, and J. Rousseau, "Volume delineation by fusion of fuzzy sets obtained from multiplanar tomographic images," *IEEE Transactions on Medical Imaging*, vol. 20, no. 12, pp. 1362–1372, 2001.

[57] A. Stefano, S. Vitabile, G. Russo, M. Ippolito, D. Sardina, M. Sabini, F. Gallivanone, I. Castiglioni, and M. Gilardi, "A graph-based method for PET image segmentation in radiotherapy planning: A pilot study," *International Conference on Image Analysis and Processing*, pp. 711–720, 2013.

[58] M. Soufi, A. Kamali-Asl, P. Geramifar, M. Abdoli, and A. Rahmim, "Combined fuzzy logic and random walker algorithm for PET image tumor delineation," *Nuclear Medicine Communications*, vol. 37, no. 2, pp. 171–181, 2016.

[59] C. Ballangan, X. Wang, M. Fulham, S. Eberl, and D. Feng, " Lung tumor segmentation in PET images using graph cuts," *Computer Methods and Programs in Biomedicine*, vol. 109, no. 3, pp. 260–268, 2013.

[60] E. Grossiord, H. Talbot, N. Passat, M. Meignan, and L. Najman, "Automated 3D lymphoma lesion segmentation from PET/CT characteristics," *IEEE International Symposium on Biomedical Imaging*, pp. 174–178, 2017.

[61] W. Ju, D. Xiang, B. Zhang, L. Wang, I. Kopriva, and X. Chen, "Random walk and graph cut for co-segmentation of lung tumor on PET-CT images," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5854–5867, 2015.

[62] P. Salembier, A. Oliveras, and L. Garrido, "Anti-extensive connected operators for image and sequence processing," *IEEE Transactions on Image Processing*, vol. 7, no. 4, pp. 555–570, 1998.

[63] L. Najman and M. Couprie, "Building the component tree in quasi-linear time," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3531–3539, 2006.

[64] P. Monasse and F. Guichard, "Scale-space from a level lines tree," *Journal of Visual Communication and Image Representation*, vol. 11, no. 2, pp. 224–236, 2000.

[65] P. Monasse and F. Guichard, "Fast computation of a contrast invariant image representation," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 860–872, 2000.

[66] V. Caselles and P. Monasse, "Geometric description of images as topographic maps," *Lecture Notes in Mathematics Series, Springer*, 2009.

[67] T. Géraud, E. Carlinet, S. Crozet, and L. Najman, "A quasi-linear algorithm to compute the tree of shapes of n-D images," *11th International Symposium on Mathematical Morphology*, vol. 7883, pp. 98–110, 2013.

[68] C. Berger, T. Géraud, R. Levillain, N. Widynski, A. Baillard, and E. Bertin, "Effective component tree computation with application to pattern recognition in astronomical imaging," *IEEE International Conference on Image Processing*, vol. 4, pp. 41–44, 2007.

[69] E. Breen and R. Jones, "Attribute openings, thinnings, and granulometries," *Computer Vision and Image Understanding*, vol. 64, no. 3, pp. 377–389, 1996.

[70] R. Jones, "Connected filtering and segmentation using component trees," *Computer Vision and Image Understanding*, vol. 75, no. 3, pp. 215–228, 1999.

[71] B. Naegel, N. Passat, N. Boch, and M. Kocher, "Segmentation using vector-attribute filters: Methodology and application to dermatological imaging," *International Symposium on Mathematical Morphology*, pp. 239–250, 2007.

[72] J. Matas, O. Chum, M. Urban, and P. T., "Robust wide baseline stereo from maximally stable extremal regions," *Image and Vision Computing*, vol. 22, no. 10, pp. 761–767, 2004.

[73] L. Grady, "Random walks for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1768–1783, 2006.

[74] S. Kakutani, "Markov processes and the Dirichlet problem," *Proceedings of the Japanese Academy*, vol. 21, pp. 227–233, 1945.

[75] M. Lev and R. Gonzalez, "CT angiography and CT perfusion imaging," *Brain Mapping: The Methods*, no. 2, pp. 427–484, 2002.

[76] T. Sørensen, "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons," *Biologiske Skrifter*, vol. 5, no. 4, pp. 1–35, 1948.

[77] R. Haralick, S. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 4, pp. 532–550, 1987.

[78] S. Hammerschlag, S. Wolpert, and B. Carter, "Computed tomography of the spinal canal," *Radiology*, vol. 121, no. 2, pp. 101–110, 1976.

[79] A. Paulino, M. Koshy, R. Howell, D. Schuster, and L. Davis, "Comparison of CT-and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer," *International Journal of Radiation Oncology, Biology, Physics*, vol. 61, no. 5, pp. 1385–1392, 2005.

[80] S. Bentzen and V. Gregoire, "Molecular imaging-based dose painting: A novel paradigm for radiation therapy prescription," *Seminars in Radiation Oncology*, vol. 21, no. 2, pp. 101–110, 2011.

[81] M. Law and B. Liu, "DICOM-RT and its utilization in radiation therapy," *RadioGraphics*, vol. 29, no. 3, pp. 655–667, 2009.

**Analyse d'images nucléaires dans une contexte multimodal et multitemporel**

Ces travaux de thèse portent sur la proposition de stratégies de segmentation des tumeurs cancéreuses dans un contexte multimodal et multitemporel. La multimodalité fait référence au couplage de données TEP / TDM pour exploiter conjointement les deux sources d'information pour améliorer les performances de la segmentation. La multitemporalité fait référence à la disposition des images acquises à différentes dates, ce qui limite une correspondance spatiale possible entre elles.

Dans une première méthode, une structure arborescente est utilisée pour traiter et pour extraire des informations afin d'alimenter une segmentation par marche aléatoire. Un ensemble d'attributs est utilisé pour caractériser les nœuds de l'arbre, puis le filtrer et projeter des informations afin de créer une image vectorielle. Un marcheur aléatoire guidé par les données vectorielles provenant de l'arbre est utilisé pour étiqueter les voxels à des fins de segmentation.

La deuxième méthode traite le problème de la multitemporalité en modifiant le paradigme de voxel à voxel par celui de nœud à nœud. Deux arbres sont alors modélisés à partir de la TEP et de la TDM avec injection de contraste pour comparer leurs nœuds par une différence entre leurs attributs et ainsi correspondre à ceux considérés comme similaires en supprimant ceux qui ne le sont pas.

Dans une troisième méthode, qui est une extension de la première, l'arbre calculé à partir de l'image est directement utilisé pour mettre en œuvre l'algorithme développé. Une structure arborescente est construite sur la TEP, puis les données TDM sont projetées sur l'arbre en tant qu'informations contextuelles. Un algorithme de stabilité de nœud est appliqué afin de détecter et d'élaguer les nœuds instables. Des graines, extraites de la TEP, sont projetées dans l'arbre pour fournir des étiquettes (pour la tumeur et le fond) à ses nœuds correspondants et les propager au sein de la hiérarchie. Les régions évaluées comme incertaines sont soumises à une méthode de marche aléatoire vectorielle pour compléter l'étiquetage de l'arbre et finaliser la segmentation.

Segmentation, TEP, TEP/TDM, arbre de formes, attributs, cancer.

**Nuclear image analysis in a multimodal and multitemporal context**

This work focuses on the proposition of cancerous tumor segmentation strategies in a multimodal and multitemporal context. Multimodal scope refers to coupling PET/CT data in order to jointly exploit both information sources with the purpose of improving segmentation performance. Multitemporal scope refers to the use of images acquired at different dates, which limits a possible spatial correspondence between them.

In a first method, a tree is used to process and extract information dedicated to feed a random walker segmentation. A set of region-based attributes is used to characterize tree nodes, filter the tree and then project data into the image space for building a vectorial image. A random walker guided by vectorial tree data on image lattice is used to label voxels for segmentation.

The second method is geared toward multitemporality problem by changing voxel-to-voxel for node-to-node paradigm. A tree structure is thus applied to model two hierarchical graphs from PET and contrast-enhanced CT, respectively, and compare attribute distances between their nodes to match those assumed similar whereas discarding the others.

In a third method, namely an extension of the first one, the tree is directly involved as the data-structure for algorithm application. A tree structure is built on the PET image, and CT data is then projected onto the tree as contextual information. A node stability algorithm is applied to detect and prune unstable attribute nodes. PET-based seeds are projected into the tree to assign node seed labels (tumor and background) and propagate them by hierarchy. The uncertain nodes, with region-based attributes as descriptors, are involved in a vectorial random walker method to complete tree labeling and build the segmentation.

Segmentation, PET, PET/CT, tree of shapes, region-based attributes, cancer.

***Discipline : INFORMATIQUE***

Université de Reims Champagne-Ardenne

CRESTIC - EA 3804

UFR Sciences Exactes et Naturelles - Moulin de la Housse - BP 1039 - 51687 Reims CEDEX 2