KIDNEY SEGMENTATION IN 4-DIMENSIONAL DYNAMIC CONTRAST-ENHANCED MR IMAGES: A PHYSIOLOGICAL APPROACH

VOREAK SUYBENG

NATIONAL UNIVERSITY OF SINGAPORE

2007

KIDNEY SEGMENTATION IN 4-DIMENSIONAL DYNAMIC CONTRAST-ENHANCED MR IMAGES:
A PHYSIOLOGICAL APPROACH

VOREAK SUYBENG

(B. ENG. SUPÉLEC)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2007

Acknowledgements

I am extremely grateful to my supervisor Associate Professor Ashraf Kassim for his guidance and concern throughout the course of this research. I am especially appreciative of his belief in this project.

I am also deeply indebted to Dr Borys Shuter for his generosity of precious advices, knowledge and experience. His invaluable help and suggestions have greatly contributed to the success of this project.

Finally, I want to express our gratitude to Professor David A. Stringer, Head of Diagnostic Imaging Department at KK Women's and Children's Hospital, for encouraging us to use his paediatric MR data sets.

Table of contents

Summary

Four-dimensional dynamic contrast enhanced-magnetic resonance imaging (4-D DCE-MRI) is an emerging clinical tool in paediatric urology. It involves administration of a contrast agent whose journey through the kidneys is monitored over time on T1-weighted images to provide detailed information on renal morphology and physiology. Such acquisitions yield a huge amount of data that makes quantification difficult. Therefore, automated image segmentation of the kidney tissues becomes important for clinical purposes.

While previous kidney segmentation methods focus on spatial information, this present thesis exploits the kinetics of contrast agent uptake in abdominal tissues. This novel approach is motivated by the reproducible pattern of time-enhancement curves which reflects known anatomy and physiology of the tissues.

Two types of techniques are proposed:

- A clustering based technique (Kohonen Neural Network, Fuzzy K-Means clustering and Expectation-Maximization clustering) identifies groups of voxels that share similar time courses and associates each group to a different tissue type. It demonstrates the ability to extract the renal parenchyma from the surrounding organs.
- A factor analysis based technique (FAMIS) with novel constraints yields factor solutions that can describe physiologically both the spatial distribution and temporal behaviour of the different tissue types. It demonstrates the ability to extract intrarenal compartments.

The proposed approach is validated over a wide range of 20 real patient MR data sets and our kidney segmentation results show good agreement with the manual segmentation performed by our expert. Especially, our FAMIS technique produces consistent results in extracting factors that are physiologically interpretable and thus, it represents a potentially useful technique in the assessment of renal function.

Index of figures

List of abbreviations

4-D DCE-MRI

Four-Dimensional Contrast Enhanced-Magnetic Resonance Imaging

ADH

Antidiuretic Hormone

CD

Collecting Duct

CT

Computed Tomography

DCT

Distal Convoluted Tubule

EM

Expectation-Maximization

FAMIS

Factor Analysis of Medical Image Series

FID

Free Induction Decay

FKM

Fuzzy K-Means

KNN

Kohonen Neural Network

Gd-DTPA

Gadolinium Diethylenetriamine Pentaacetic Acid or Gadopentetate Dimeglumine

IVU

Intravenous Urography

LOH

Loop of Henle

MIP

Maximum Intensity Projection

MRI

Magnetic Resonance Imaging

MRU

Magnetic Resonance Urography

PCT

Proximal Convoluted Tubule

RF pulse

Radio Frequency pulse

ROI

Region of Interest

Tc-MAG3

Technetium Mercaptoacetyltriglycine

TE

Echo Time

TEC

Time-Enhancement Curve

TIC

Time-Intensity Curve

TR

Repetition Time

US

Ultrasound

USPIO

Ultrasmall Superparamagnetic Iron Oxide

VOI

Volume of Interest

1 Introduction

1.1 Background on imaging modalities

Magnetic Resonance Imaging (MRI) has revolutionized the diagnosis of many types of illness. This progress was possible due to the concomitant development and use of contrast agents which accumulate in specific tissues or organs to improve the diagnosis of these regions.

Recent investigations indeed recognize that MR images collected over time by dynamic contrast-enhanced MRI (DCE-MRI) have the potential to provide both anatomic and physiologic information: the contrast agents are administered not only to enhance the visibility of an organ of interest, but also to analyse its biological function by measuring the temporal change of the spatial distribution of the contrast agent.

In paediatric renal imaging, the use of gadolinium-enhanced excretory MR imaging of the kidney, also called dynamic contrast-enhanced MR Urography (DCE-MRU), is considered as a major breakthrough. The contrast agent, commonly Gd-DTPA, is designed to concentrate mostly at the kidney which appears brighter. Localization and differentiation of this organ may then be easier and information about its concentration and excretory function may also be analysed by visualizing the passage and elimination of the contrast agent [62] [64].

Despite their widespread use, other imaging modalities like nuclear scintigraphy, Doppler ultrasound (US), intravenous urography (IVU) and computed tomography (CT) have inherent shortcomings:

- Routinely, nuclear scintigraphy, US and IVU provide only 2-D images of the kidney.
- Nuclear scintigraphy provides functional information but relatively poor spatial resolution.
- US has better spatial resolution but little functional information.
- IVU and CT have good spatial resolution and the potential to provide information on renal excretion by demonstrating the course of a bolus of contrast agent through the glomeruli but subject paediatric patients to nephrotoxic contrast agents.

- Nuclear scintigraphy, IVU and CT use ionizing radiation that has potential deleterious effects.

On the contrary, 4-D DCE-MRU (3-D + time) can provide in a single study both very good spatial and adequate temporal resolution for functional assessment without the use of nephrotoxic contrast agents and ionizing radiation. Although expensive and not always available, it has several advantages in the examinations of the kidneys because it supplies detailed information regarding anatomy, function and drainage and thus, provides some completely new insights into renal pathophysiology [24] [50] [63] [64] [66], renal transplant functional assessment [52] [57] and functional renal volume [43] [61]. Therefore, it is a promising tool for the assessment of the urinary system and with future advances in automated image analysis, it could emerge as the functional renal imaging modality of choice.

1.2 Motivations

4-D DCE-MRU examinations always include an image processing routine to extract the different physiological compartments of the kidneys from the surrounding tissues. Currently, this may be done by comparing, on a slice-by-slice basis, volumetric images of a same patient acquired over the sequence after manually placing volumes of interest (VOI) that enclose the renal cortex, the renal medulla and the renal pelvis. Then, the mean signal intensity within each VOI at each time point is computed and plotted against time, resulting in a so-called renogram describing renal function and urinary excretion.

But such a manual technique in 4-D DCE-MRU is repetitive, laborious and time consuming because of the large amount of data: a patient acquisition can easily generate a sequence of more than 200 volumes, each of 12-20 slices. Moreover, it is highly dependent on user-defined volumes of interest, especially in poorly functioning kidneys, making the results subject to large inter-observer and intra-observer variability.

In this context therefore, automatic segmentations that can extract the kidneys in 3-D are becoming an increasingly important image processing step for clinical diagnosis. These tasks are especially challenging in DCE-MRI because:

- Image intensity values change rapidly over the time course reflecting the rapid wash-in and wash-out of contrast agent from the abdominal tissues.
- Contrast agent uptake in poorly functioning kidneys may be reduced, resulting in weak boundaries between different tissues or disjoined bright regions.

- Kidneys are located in a crowded soft-tissue environment and contrast agent also enhances the neighbouring tissues such as spleen, liver, intestine and aorta.
- The image quality is subject to the amount of noise, motion artefacts and the current technological limitations of the imaging modality.

In this thesis, novel techniques for kidney segmentation in 4-D DCE-MRI based on a physiological and temporal approach are presented. They are motivated by the observation that a reproducible pattern of time-enhancement curves characterises each abdominal organ. The methods are tested on 20 real patient MR data sets and the results show consistent agreement with the segmentations provided by the domain expert.

The algorithms are implemented in MATLAB(R) 7.0, The MathWorks, Inc., to facilitate integration into an existing GUI (Graphical User Interface) used by doctors of the National University Hospital (NUH).

A brief description of our 20 MR data sets are given in **Appendix**.

1.3 Contribution summary

The key contributions of this work are summarized below:

- A novel 4-D DCE-MRI approach which analyses the kinetics of contrast agent uptake in abdominal tissues to segment kidneys is proposed. It is motivated by the reproducible pattern of time-enhancement curves that can be explained by known anatomy and physiology.
- Clustering based methods are applied to regroup voxels that share a similar time course and a factor analysis based method with new constraints is developed to segment not only kidneys from surrounding tissues but also intrarenal compartments.
- The performance of the techniques is demonstrated on 20 real MR patient data sets and it is shown that the factor analysis based method yields promising results for a clinical routine.

1.4 Organization of the thesis

The contents of the key chapters that follow in the remainder of this thesis are described briefly below:

- **Chapter 2: Dynamic contrast-enhanced MR Urography.** Some of the necessary concepts to comprehend the subtleties of MR imaging are reviewed. Insights into the anatomy and physiology of the kidneys are especially given to demonstrate the renal enhancement pattern observed in DCE-MRU.
- **Chapter 3: Related works in kidney segmentation.** A literature review of medical image segmentation of the kidneys is presented.
- **Chapter 4: Proposed methods for kidney segmentation.** The principles and the algorithms of the proposed methods based on cluster analysis and factor analysis to segment kidney tissues in DCE-MRU are described in detail.
- **Chapter 5: Experimental results and Discussion.** The performance of the segmentation results is evaluated and discussed.
- **Chapter 6: Conclusion.** The results are summarized and future research directions are indicated.

2 Dynamic contrast-enhanced MR Urography

A detailed explanation of MRI is beyond the scope of this thesis and the reader is invited to refer to [31], [34] or [69] for more detailed information.

2.1 Basic MRI physics

2.1.1 Macroscopic magnetization

Hydrogen nuclei exhibit a magnetic moment due to the spinning of their proton around their own axis. When the hydrogen protons are placed within a powerful external magnetic field of strength B_0, they align themselves mostly toward the direction of the magnetic field (spin-up or parallel) and experience a precession by rotating around the direction of the field (Figure 1). The precessional frequency is proportional to the external magnetic field and is given by the Larmor equation:

$$\omega_0 = \gamma B_0 \tag{2.1.1}$$

- ω_0 is the angular frequency of precession in radians and is called the Larmor frequency.
- γ is a constant called gyromagnetic ratio which is a distinguishing characteristic of each element. As hydrogen is naturally abundant in human tissues which are made of water at 75%, clinical MRI is confined to imaging the hydrogen nucleus for which $\gamma = 42.57\,\mathrm{MHz/T}$.
- B_0 is the magnetic field strength in Tesla.

Consequently, within a magnetic field and at equilibrium, the vector sum of all the magnetic moment vectors of the individual protons, called net magnetization or macroscopic magnetization $\boldsymbol{M} = \boldsymbol{M_0}$, can be described by a longitudinal component $\boldsymbol{M}_z = \boldsymbol{M} = \boldsymbol{M_0}$ aligned with the field, along the z axis, while the transverse component $\boldsymbol{M}_{xy}$, lying on the x-y plane (also known as the transverse plane), is null because the individual protons do not precess in phase with each other (Figure 2).

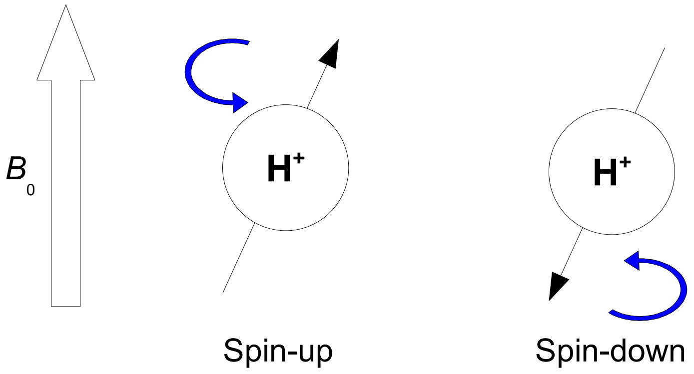


Figure 1: Precession of hydrogen nuclei

Hydrogen nuclei exhibit a magnetic moment (straight black arrows). Within a magnetic field B_0, they precess around the direction of the magnetic field (as indicated by the curved blue arrows) and are aligned mostly spin-up or parallel with the magnetic field.

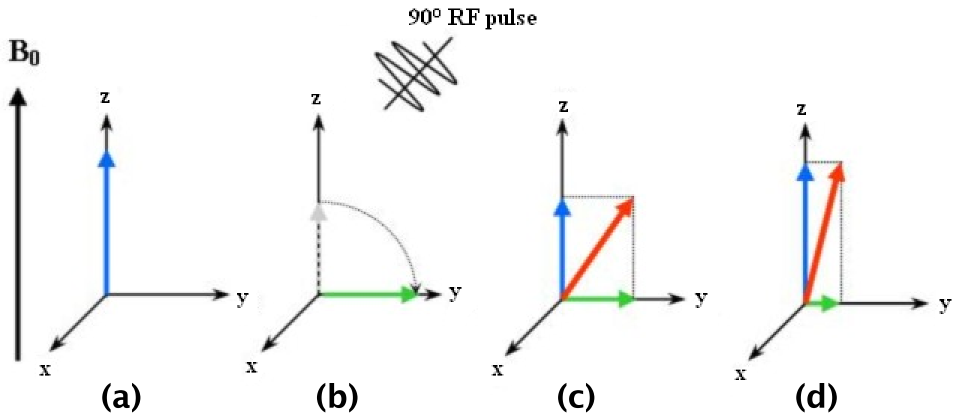


Figure 2: Effect of a 90° RF pulse on the macroscopic magnetization

- **(a)** ***Equilibrium:*** *within a magnetic field B_0 and at equilibrium $\mathbf{M_0}$, the macroscopic magnetization $\mathbf{M=M_0}$ is described by a longitudinal component $\mathbf{M_z}$ (blue arrow) while the transverse component $\mathbf{M_{xy}}$ is null: $\mathbf{M=M_0=M_z}$.*
- **(b)** ***Magnetic resonance and excitation:*** *a 90° RF pulse at the Larmor frequency perturbs the equilibrium: it tips down the macroscopic magnetization **M** at 90° so that the longitudinal magnetization $\mathbf{M_z}$ disappears while the transverse magnetization $\mathbf{M_{xy}}$ (green arrow) appears: $\mathbf{M=M_{xy}}$.*
- **(c) (d)** ***Relaxation:*** *when the RF pulse is turned off, the macroscopic magnetization **M** (red arrow) returns back to the equilibrium $\mathbf{M_0}$. The longitudinal relaxation refers to the recovery of $\mathbf{M_z}$ back to $\mathbf{M_0}$ and the transverse relaxation refers to the decay of $\mathbf{M_{xy}}$ back to the null vector. Notice that the transverse relaxation is always faster than the longitudinal relaxation (T2<T1).*

http://upload.wikimedia.org/wikipedia/commons/2/22/90pulse.jpg

The original image is under the GNU Free Documentation License and was modified to fit the explanations given in this thesis.

2.1.2 Magnetic resonance and excitation

Magnetic resonance is accomplished by applying a radio frequency (RF) pulse at exactly the same frequency as the Larmor frequency that perturbs the proton equilibrium: the net magnetization vector $\boldsymbol{M}$ moves away from alignment in the z direction and rotates towards the x-y plane so that the longitudinal magnetization $\boldsymbol{M}_z$ decreases whereas the transverse magnetization $\boldsymbol{M}_{xy}$ appears (Figure 2). The angle of rotation α is called the tip angle or flip angle and is proportional to the strength B_1 and the duration t_p of the RF pulse according to the following equation:

$$\alpha = \gamma B_1 t_p \tag{2.1.2}$$

2.1.3 Relaxation times

Relaxation refers to the return to equilibrium of the net magnetization $\boldsymbol{M}$ back to $\boldsymbol{M}_0$ as soon as the RF pulse is turned off (Figure 2). It combines 2 different mechanisms:

- **Longitudinal relaxation** or **T1 relaxation** (Figure 3) corresponds to the recovery of the longitudinal magnetization $\boldsymbol{M}_z$ back to $\boldsymbol{M}_0$ according to an exponential curve characterised by the tissue-specific time constant T1. It is also known as spin-lattice relaxation because the recovery is due to the release by the excited protons of RF energy back to the surrounding lattice. After time T1, the longitudinal magnetization $\boldsymbol{M}_z$ has returned to 63% of its value $\boldsymbol{M}_0$ at equilibrium.
- **Transverse relaxation** or **T2 relaxation** (Figure 4) corresponds to the decay of the transverse magnetization $\boldsymbol{M}_{xy}$ to a null vector according to an exponential curve also characterised by the tissue-specific time constant T2<T1. It is also known as the spin-spin relaxation time. Indeed, the RF pulse at the Larmor frequency forces the individual protons to precess in phase but when it is turned off, this phase coherence is progressively lost due to the interaction between excited protons which modify their precession rate. After time T2, the transverse magnetization $\boldsymbol{M}_{xy}$ has dropped to 37% of its initial magnitude considered at the moment when the RF pulse was turned off.

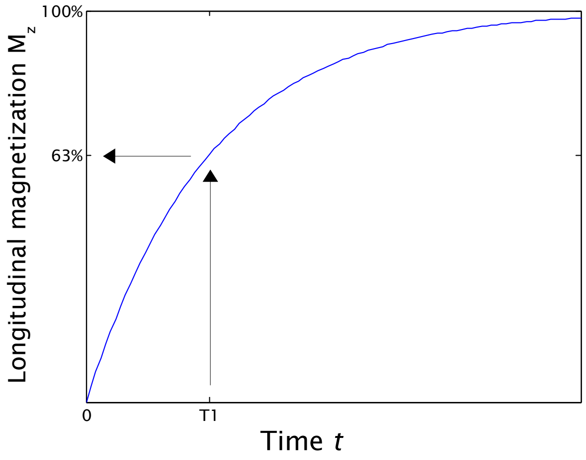


Figure 3: Longitudinal relaxation T1

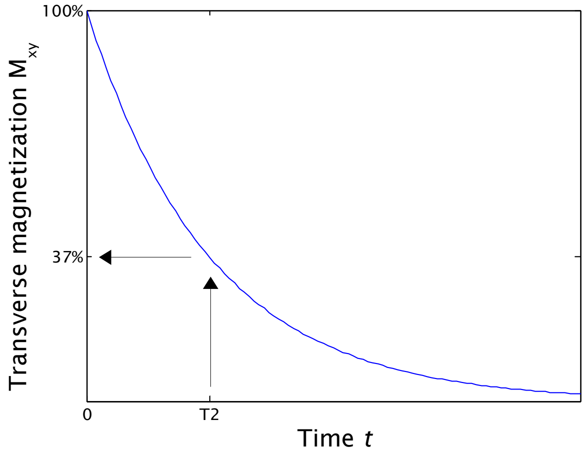


Figure 4: Transverse relaxation T2

2.2 Magnetic resonance imaging

2.2.1 Tissue contrast

Basically, the MR signal and thus, the contrast in MR images is produced by weighting (enhancing or reducing) the differences in proton density (PD) and relaxation times (T1 and T2) that characterize each tissue (Figure 5). This is done by using different pulse sequences (spin echo sequence in section **2.2.3.** or gradient echo sequence in section **2.2.4.**) and by tuning their parameters. Depending on the tissue characteristic that is emphasized, three types of image contrast can be produced:

- PD-weighted images where tissues with high proton densities give generally the highest signal intensities and thus the brightest voxels.
- T1-weighted images where tissues with the shortest T1 give the highest signal intensities and thus the brightest voxels.
- T2-weighted images where tissues with the longest T2 give the highest signal intensities and thus the brightest voxels.

As it shall be explained, a spin echo sequence and a echo gradient sequence can both produce PD-, T1-, T2-weighted images depending on the choice of their parameters.

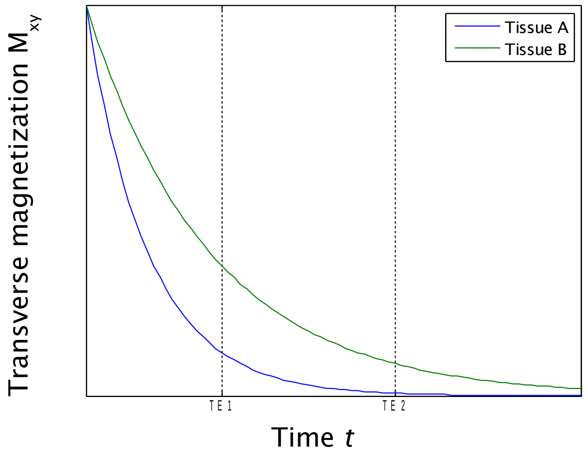


Figure 5: Development of tissue contrast

- *The figure considers the transverse magnetization decay in T2 of two different tissues.*
- *t=TE1 provides better tissue contrast than t=TE2.*

2.2.2 MR signal and Free Induction Decay (FID)

The MR signal that can be measured by the RF coils is always the current signal induced by the variations of the transverse magnetization vector $\boldsymbol{M}_{xy}$. This signal is oscillating at Larmor frequency and its envelope decays exponentially to zero at a time constant T2* faster than T2 due to both the loss of phase coherence between individual protons and the inhomogeneities in the external magnetic fields. When it is measured after transmission of a RF pulse at the Larmor frequency, it is known as the Free Induction Decay (FID) (Figure 6). But actually, spin echo or gradient echo signals rather than FID are measured in MR imaging.

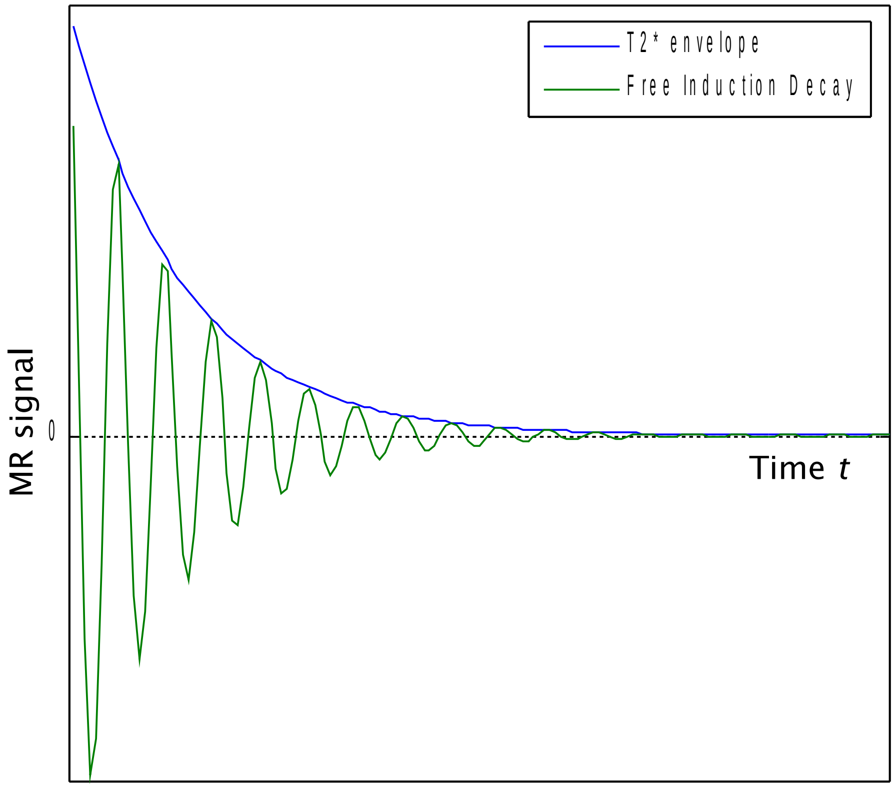


Figure 6: A free induction decay with a T2* envelope

2.2.3 Spin echo sequence

Pulse sequence

The spin echo sequence is based on the repetition of 90° and 180° RF pulses (Figure 7):

- A 90° RF pulse at the Larmor frequency tips the net magnetization $\boldsymbol{M}$ down at 90° so that:
  - $\boldsymbol{M}$ lies in the $x\text{-}y$ plane.
  - Longitudinal magnetization $\boldsymbol{M}_z$ disappears.
  - Transverse magnetization appears and $\boldsymbol{M}_{xy}=\boldsymbol{M}_0$ in amplitude (due to complete phase coherence between individual protons).
- Once the 90° RF pulse is turned off:
  - Longitudinal magnetization $\boldsymbol{M}_z$ recovers with time-constant T1.
  - Transverse magnetization $\boldsymbol{M}_{xy}$ decays with time-constant T2* due to both the loss of phase coherence between individual protons (random) and the inhomogeneities in magnetic fields (non-random).
- A 180° RF pulse at the Larmor frequency applied at time TE/2 after the 90° RF pulse reverses the non-random dephasing so that protons come back into phase to increase the transverse magnetization $\boldsymbol{M}_{xy}$ and form, at time TE (Echo Time), the spin echo which is the sampled MR signal. However, due to decay in T2, the signal of the spin echo is not as high as the initial transverse magnetization.
- Then, the protons dephase again and the transverse magnetization $\boldsymbol{M}_{xy}$ decays due to T2* effect. This sequence of 90° and 180° RF pulses is repeated multiple times again depending on the MR image size.

The spin echo sequence has two parameters:

- The echo time, TE, is the time between the 90° RF pulse and the maximum of the spin echo, corresponding to the MR signal sampling. TE/2 is the time between the 90° RF pulse and the 180° RF pulse.
- The repetition time, TR, is the time between two consecutive 90° RF pulses.

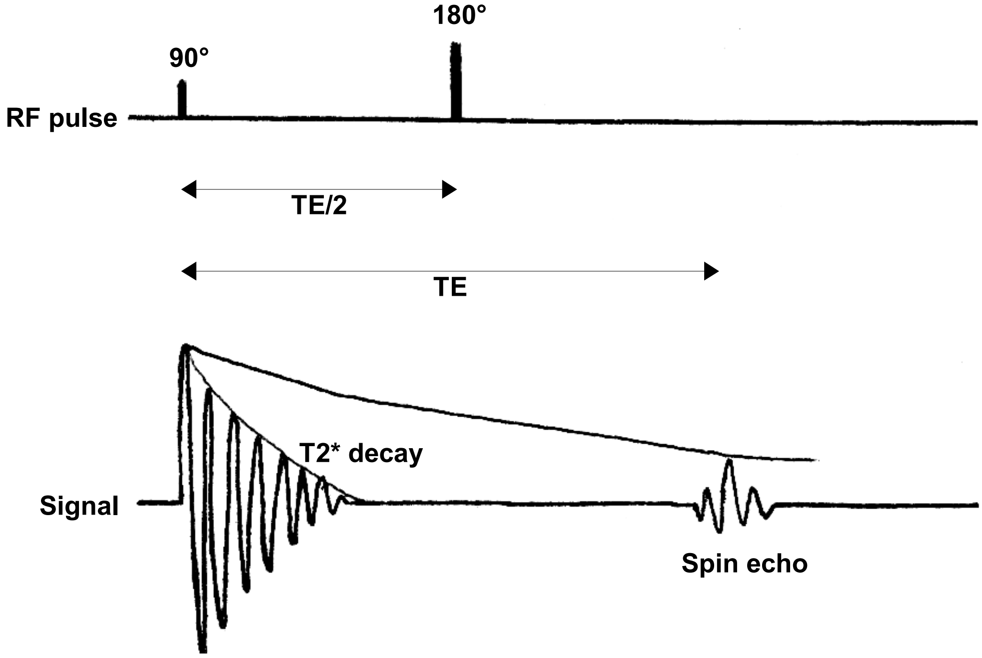


Figure 7: Formation of a spin echo at time TE

Development of tissue contrast

A T1-weighted image is taken as example to explain how the tissue contrast is achieved in a spin echo sequence. In a T1-weighted image, the contribution due to T1 must be emphasized while those of T2 and PD must be reduced. This is done by using a short TR and a short TE:

- Each tissue is characterised by a relaxation time T1 for the recovery of their longitudinal magnetization M_z. So, a short TR enhances these differences in longitudinal recovery between tissues with different T1.
- A short TE ensures that protons are not given sufficient time to dephase and thus that the transverse magnetization differences between tissues due to decay in T2 have a minimum effect on the differences due to the characteristic T1 of the different tissues.

2.2.4 Gradient echo sequence

Pulse sequence

There are two main differences that distinguish a gradient echo sequence from a spin echo sequence (Figure 8):

- The exciting RF pulse yields a flip angle lower than 90°. Consequently, the longitudinal magnetization $\boldsymbol{M}_z$ flips incompletely into the $x\text{-}y$ plane so that a transverse magnetization $\boldsymbol{M}_{xy}$ appears while a certain amount of longitudinal magnetization $\boldsymbol{M}_z$ still remains.
- Once the RF pulse is turned off, $\boldsymbol{M}_z$ recovers with time constant T1 while $\boldsymbol{M}_{xy}$ decays with time constant T2* like in the spin echo sequence.
- But then, instead of using a 180° RF pulse, the gradient echo sequence rather uses a magnetic field gradient reversal to rephase the protons:
    - A negative magnetic field gradient $-G_x$ is first applied, forcing the protons to dephase much faster than in T2*. That is, depending on its position respectively to the gradient, each proton is forced to precess faster or slower, until the transverse magnetization $\boldsymbol{M}_{xy}$ is eliminated.
    - Then, the magnetic field gradient is reversed to yield a positive amplitude. So, protons which were previously precessing faster start to slow down, while those which were precessing slower start to speed up. As a result, all the protons eventually come back into phase to restore the transverse magnetization $\boldsymbol{M}_{xy}$ and form, at time TE (Echo Time), the gradient echo, which is the sampled MR signal. However, the magnetic field gradient reversal does not compensate for the dephasing due to T2* and thus, the height of the gradient echo signal decays in T2*. For this reason, a gradient echo should rather be considered as a FID rather than a true echo and it produces what it should be called T2*-weighted images rather than T2-weighted images.

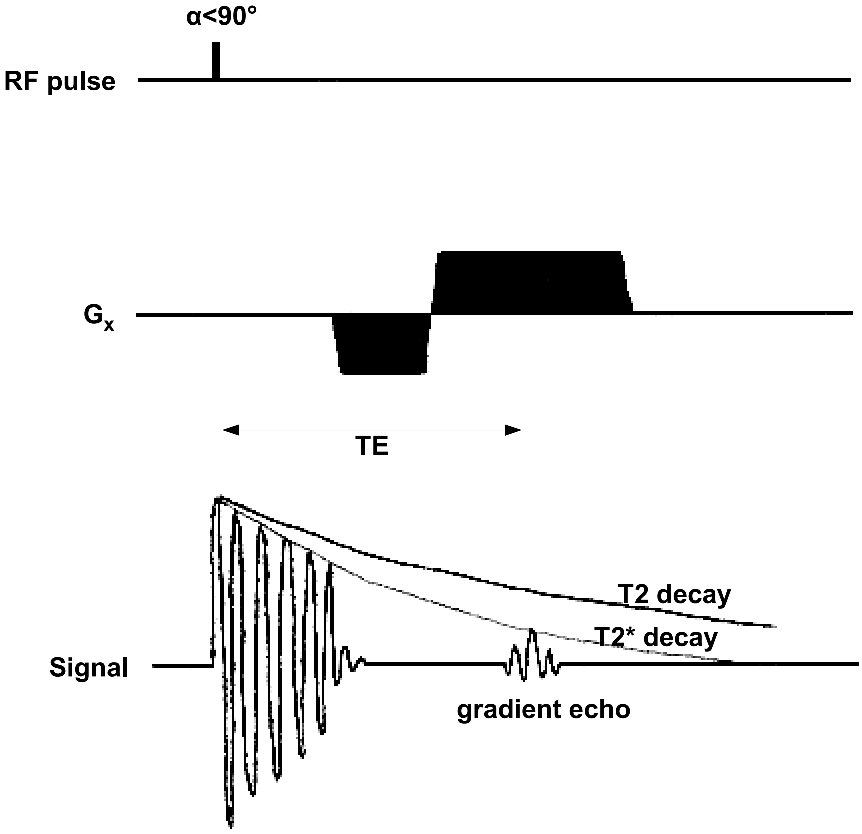


Figure 8: Formation of a gradient echo at time TE

The gradient echo sequence has three parameters:

- The echo time, TE, which is the time between the RF pulse and the maximum of the gradient echo, corresponding to the MR signal sampling.
- The repetition time, TR, which is the time between two consecutive RF pulses.
- The flip angle, α , which is caused by the RF pulse.

Advantage: faster acquisitions

The main advantage of a gradient echo sequence over a spin echo sequence is its reduced scan time by shortening the TR. This is because a partial flip angle determines both the amount of longitudinal magnetization remained in the z axis and that tipped into the x-y plane:

- With a 90° flip angle, $\boldsymbol{M}_z$ becomes null. So, if a small TR is used to speed up the scan, $\boldsymbol{M}_z$ has not sufficient time to recover. Thus, it decreases progressively with each pulse and hence, it also yields subsequent decreased $\boldsymbol{M}_{xy}$ and signals.
- But with a partial flip angle, say 30°, a residual $\boldsymbol{M}_z$ persists. Thus, its recovery can be more complete. So, even if TR is shortened to reduce the scan time, $\boldsymbol{M}_z$ can still keep an adequate amplitude after each pulse to tip a transverse magnetization $\boldsymbol{M}_{xy}$ into the x-y plane and thus, induce a signal.

Steady-state transverse magnetization

Thus; in a gradient echo sequence, the scan time can be reduced by shortening the TR. But if the TR used is shorter than the T2* that characterises the tissues, then the protons are not given sufficient time to dephase (that is, the transverse magnetization has not enough time to decay completely). As a result, there is a permanent residual transverse magnetization that remains at the end of each cycle, just before the next RF pulse, and eventually, after a few cycles, it reaches a steady state value. This steady state transverse magnetization, denoted $\boldsymbol{M}_{ss}$, may affect the tissue contrast, producing more T2*-weighting in MR images. Therefore, in spoiled gradient echo sequences, it is possible to reduce T2*-weighted images with a gradient spoiler which cancels out any residual transverse magnetization prior to each RF pulse by enhancing the proton dephasing.

Development of tissue contrast

A gradient echo sequence also produces new contrasts between tissues because there are three parameters that can be tuned by the radiologist: TE, TR and the flip angle α. As a result, it gives rise to several mixtures of PD-, T1-, T2*-weighting images:

- A short TE reduces T2* weighting like in a spin echo sequence.

- A short TR enhances T2* weighting. Indeed, if TR is smaller than T2*, then a steady state transverse magnetization M_{ss} appears and contributes to the signal.
- A small flip angle α reduces T1 weighting. Indeed, for small flip angles, a longitudinal magnetization M_z persists and thus, it can recover completely faster. So, the differences in longitudinal recovery between the tissues will be minimized.

2.2.5 Contrast agent in MR urography

The use of contrast agent in MR urography, often designated as dynamic contrast enhanced MRU (DCE-MRU), is motivated by the necessity for clinical purposes to obtain both morphological and functional information of the urinary tract:

- The use of contrast agent enhances the contrast between different tissues and may also enhance the contrast difference between the renal pathological lesions and the surrounding unaffected renal tissues.
- It offers information about the excretory function of the kidneys since the passage of contrast agent through the urinary tract can be visualized.

The contrast agent, commonly Gd-DTPA (gadolinium diethylenetriamine pentaacetic acid or gadopentetate dimeglumine), is injected into the patient intravenously. The paramagnetic properties of Gd-DTPA are explained by the gadolinium ion while the physiologic properties are explained by the attached molecule DTPA:

- Gadolinium decreases the relaxation times T1 and T2 of hydrogen protons. Hence, on T1-weighted images, it enhances the signal intensity of tissues where it is absorbed. It is a toxic element but when chelated to DTPA, it becomes safe while its paramagnetic properties are still preserved.
- DTPA is a substance that is excreted exclusively in the kidneys by glomerular filtration with no tubular secretion or reabsorption.

So, when bound together, the path of Gd-DTPA through distinct regions of the kidneys (cortical vasculature, medullary tubules and pelvis) can be monitored over time using T1-weighted spin echo imaging or T1-weighted gradient echo imaging to obtain information about renal perfusion, uptake and excretion.

Gd-DTPA has the same properties as Tc-DTPA (technetium-DTPA) used for renal scintigraphy and a good correlation can be obtained between the contrast agent elimination assessed by DCE-MRU and the conventional radionuclide methods. Some groups ([23] [51]) even show that DCE-MRU can be a good alternative to conventional radionuclide methods.

It must be mentioned that several other MR contrast agents [68] are widely available and in broad clinical use, but only Gd-DTPA will be referred in the remainder of this thesis.

2.3 MR urography imaging of renal function

2.3.1 Physiologic background of the kidneys

Introduction

In the first pass, 20% of the injected dose of contrast agent reaches the kidneys [54]. But this is only true in approximately the first 20 seconds following injection. After that, there is considerable dilution, recirculation and re-enhancement of the kidney.

The blood flow (Figure 9) (Figure 10) (Figure 11) comes from the **renal arteries** which branch from the abdominal **aorta**. Upon entering the **renal hilus**, it is delivered to the nephrons via **interlobular arteries** that separate many times into **afferent arterioles**.

The **nephron** (Figure 12) is the basic functional unit of the kidney. There are more than a million in each kidney. Their role is to clean the blood supply by:

- filtration in the **glomerulus**
- reabsorption of water in the **tubules**
- secretion of substances in the **tubules**

Blood cleaned by the nephrons returns back to the general blood circulation via the **renal veins** with a reduced amount of key solutes. Those key solutes, which include toxic and waste materials like contrast agents, are concentrated in the nephrons, subsequently sent in the form of urine into the **bladder** via the ureters, and finally excreted into the outside environment via the **urethra**.

To follow the path of Gd-DTPA through the kidneys (Figure 9) (Figure 10) (Figure 11) (Figure 12) and understand its changes of concentration, the function of the nephron must be explained [24].

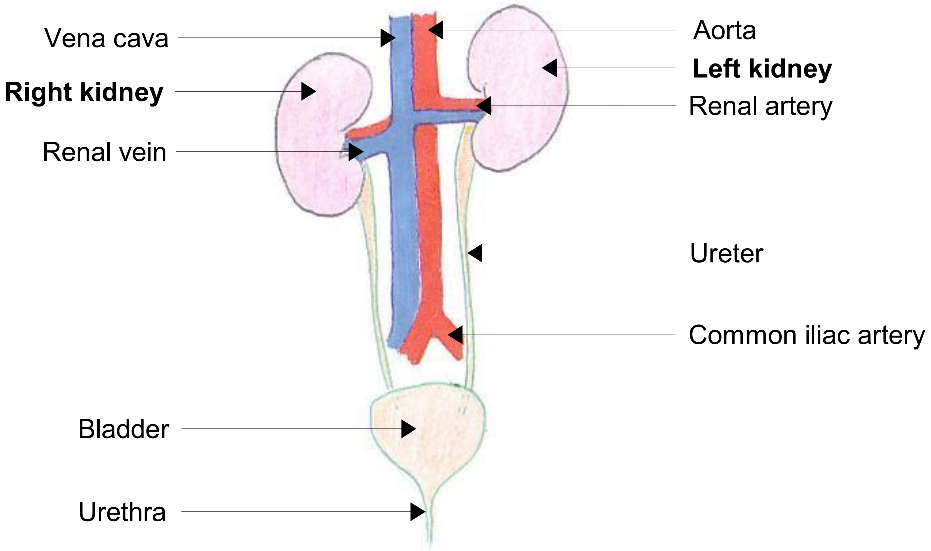


Figure 9: Diagram of the urinary system

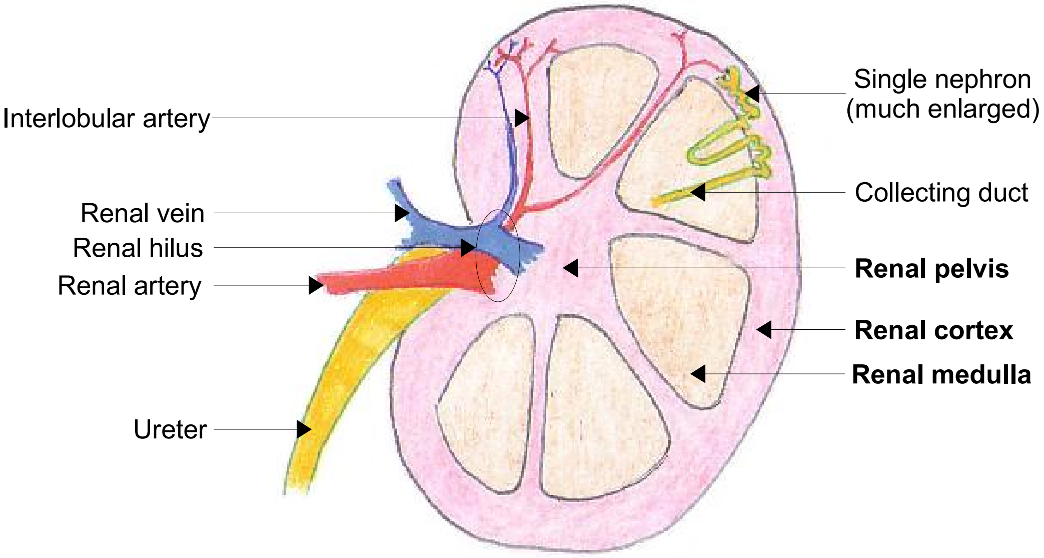


Figure 10: Gross anatomy of the left kidney

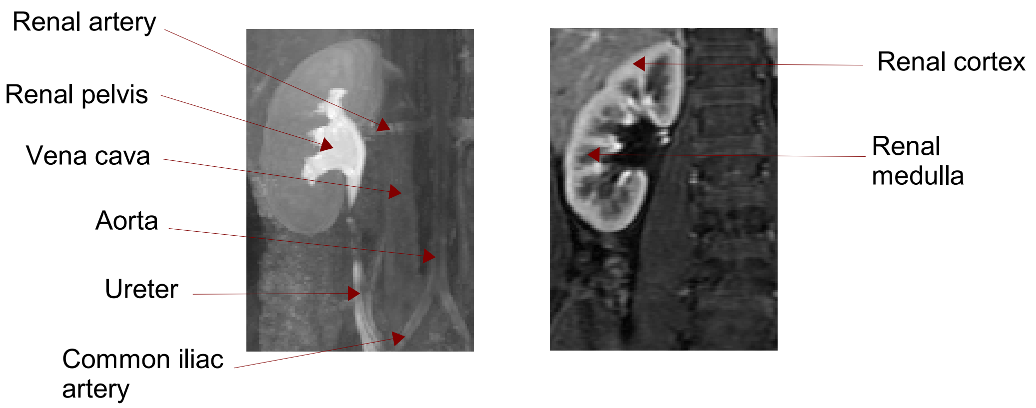


Figure 11: Right kidney on a MR image

- *Left: right kidney on a maximum intensity projection (MIP) derived from the full volume (16 slices) and the full dynamic sequence (24 time points).*
- *Right: right kidney at t=5.7min, during vascular phase, slice 8*
- *Data set 040205bpa1.*

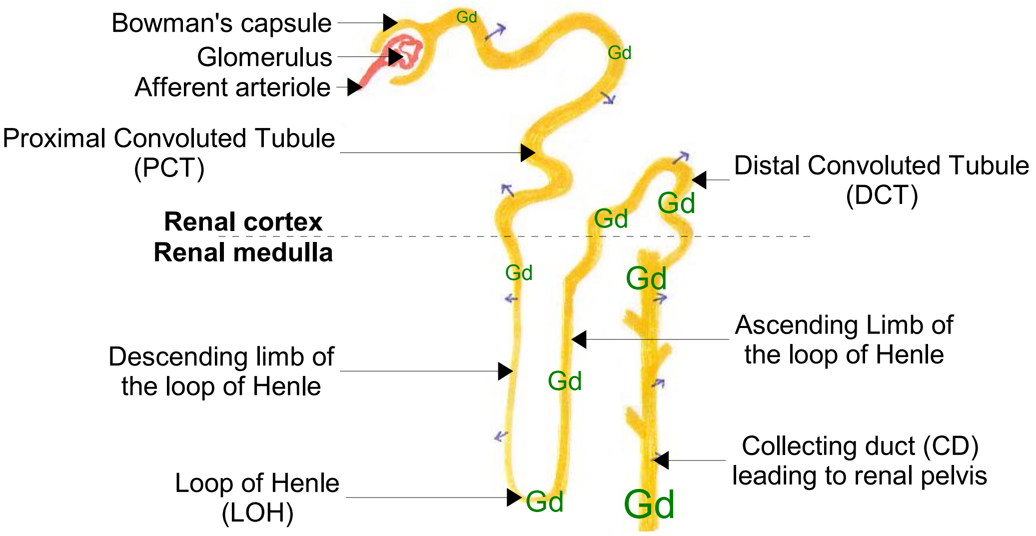


Figure 12: Journey of Gd-DTPA through the kidney nephron

- *Gd-DTPA goes down from the kidney cortex to the medulla via the descending limb of the loop of Henle.*
- *Then, it goes back to the cortex via the ascending limb of the loop of Henle and the DCT. Finally, it returns to the medulla again and reaches the renal pelvis in the form of urine via the CD.*
- *Blue arrows indicate water reabsorption and enlarging Gd indicate changes in the intratubular concentration of Gd-DTPA from PCT to CD.*
- *This pathway and concentration mechanism is essentially identical for all Gadolinium, iodine-based CT contrast agents, Tc-DTPA, creatinine and many other molecules.*

Glomerular filtration

In a single pass, most of the water and around 20% of the Gd-DTPA delivered from the blood flow are forced by high pressure out of the **glomerular capillaries** to pass into the **Bowman's capsule**. They compose the so-called primary urine or glomerular filtrate. The remaining 80% of the Gd-DTPA leaves the kidney and returns to the general blood circulation via the **renal veins** and the ascending **vena cava**.

Tubular reabsorption

After glomerular filtration, the glomerular filtrate enters the **proximal convoluted tubules** (PCT) where the concentration of Gd-DTPA is briefly diluted during the first pass by the existing fluid. However, at the end of the PCT, most of the water composing the glomerular filtrate is absorbed by osmosis so that only one third of the filtrate volume enters the outer medulla.

In the **medulla,** the epithelium covering the **descending limb of the loop of Henle** is permeable to water and thus, absorbs the remaining water in the tubular fluid which progressively becomes more concentrated in Gd-DTPA. But then, unlike the descending limb, the **ascending limb of the loop of Henle** and earlier sections of the **distal convoluted tubules** (DCT) are impermeable to water but permeable to solutes and urea. So, on the way back to the cortex, the tubular fluid is a bit diluted. This results in a so-called counter-current mechanism that yields the highest concentration of Gd-DTPA in the top of the **loop of Henle** (LOH), that is, in the inner part of the medulla.

Finally, depending on the release of antidiuretic hormone (ADH), more or less water is reabsorbed again in the latter section of the **distal convoluted tubules** and in the **collecting ducts** (CD). The definitive fluid, now called urine, reaches a maximal concentration in the **renal pelvis** and is drained into the **bladder** via the **ureters**.

Tubular reabsorption explains that during the passage of the glomerular filtrate through the tubules and the collecting ducts, most of the nutrients are sent back to the blood stream and finally only 1% of the glomerular filtrate actually leaves the organism.

Tubular secretion

Tubular secretion is not relevant to Gd-DTPA, but for information only, it is a process that keeps the blood pH at a normal level by especially adding ions in the tubular fluid.

2.3.2 Renal enhancement pattern

If Gd-DTPA is intra-venously injected and if the signal changes related to the perfusion are tracked over time for abdominal organs, the signal intensity of renal cortex, renal medulla, spleen and liver will all demonstrate first a peak enhancement and subsequently show a continuous decline reflecting the contrast agent uptake from the blood and its later expulsion.

But although the appearance of the excreted contrast agent through the kidney depends on the pulse sequence that is used and on the concentration of Gd-DTPA that is administered, there is a reproducible enhancement pattern of signal intensity changes described by three typical phases (Figure 13) (Figure 14) [40] [59] which can be explained by known anatomic and physiologic characteristics of the organs:

- A **vascular phase** occurs immediately after the injection of the contrast agent which reflects the contrast agent uptake from the blood.
    - This phase can be identified by the enhancement of the renal cortex whose signal intensity describes a steep linear rise until a peak is reached while the medulla is still hypointense. It is explained by the arrival at plasma concentration of the first bolus of Gd-DTPA into the kidney via the cortical vasculature and glomerular capillaries.
    - The spleen exhibits an inhomogeneous enhancement with patchy areas of alternating higher and lower signal intensity that some authors relate to the varying blood flow rate through the cords of red pulp [4] [11] [12] [14] [20].
    - The liver enhances slightly later [12] which reflects its dual blood supply in two times. Indeed, approximately 25% of the blood flow to the liver first comes from the hepatic artery and then approximately 75% from the portal venous system which brings nutrient-rich blood from the spleen and other abdominal organs. Consequently, the initial uptake of Gd-DTPA in the liver can be described in two phases: a first subtle arterial phase which corresponds to the arrival of a small quantity of Gd-DTPA and thus to a minimal liver enhancement; and a second subsequent portal phase which corresponds to the arrival of venous Gd-DTPA and thus a slightly later liver enhancement peak.
    - The aorta, which supplies the blood to the abdominal organs, and its bifurcation into common iliac arteries (future femoral arteries) also enhance and may be visible in some slices.

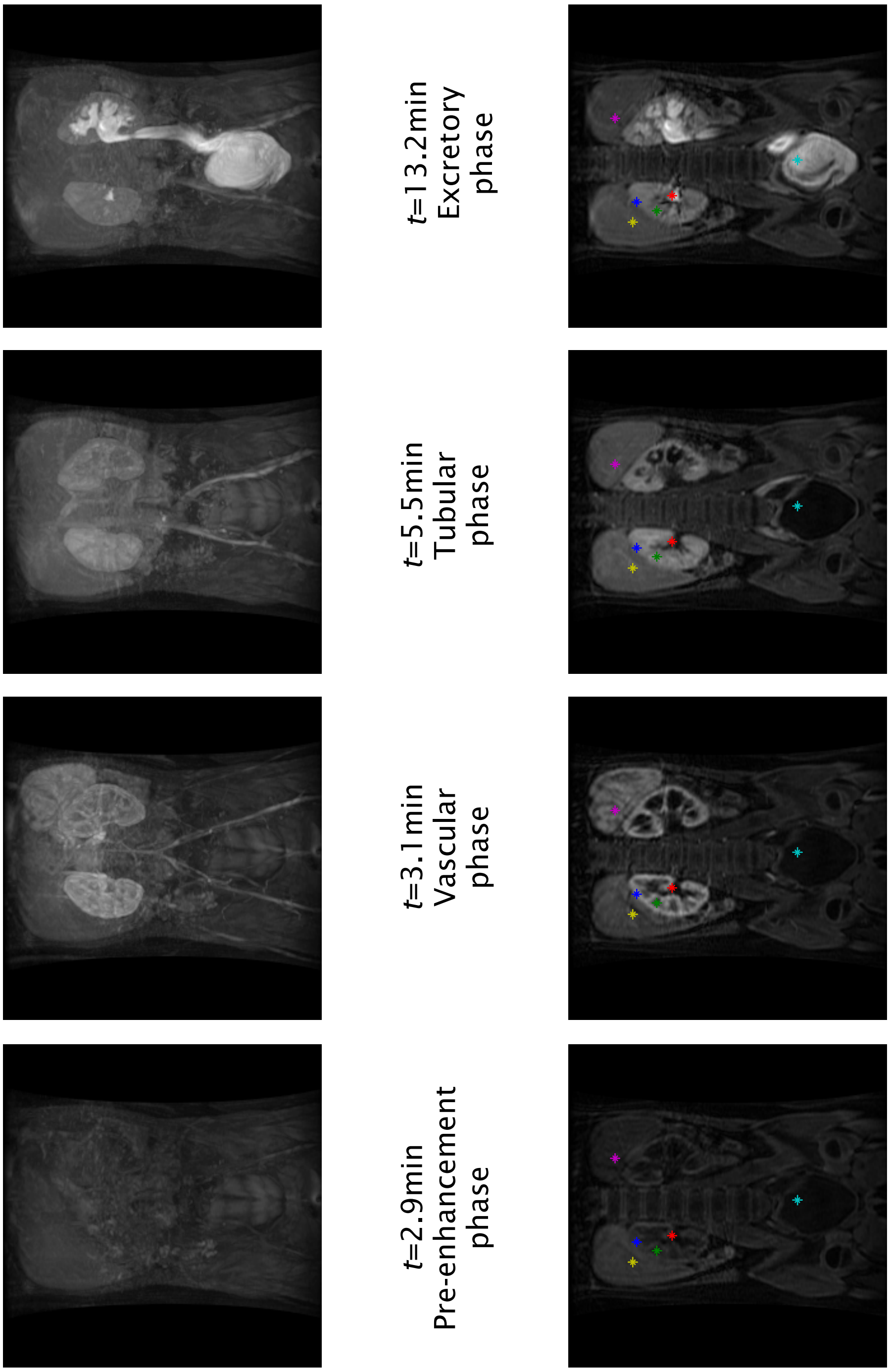


Figure 13: Selected DCE-MRU images

- *Left column shows maximum intensity projection (MIP) derived from four image volumes (14 slices) considered respectively at four separate time points (t=2.9min, t=3.1min, t=5.5min, t=13.2min).*
- *Right column shows slice 6 of the four image volumes considered respectively at the four separate time points (t=2.9min, t=3.1min, t=5.5min, t=13.2min). Voxels representative of the different physiological tissue types (cortex, medulla, pelvis, bladder, spleen and liver) are selected (cf crosses) and their signal intensities are averaged with those of their 8 neighbours and plotted over time in Figure 14.*
- *Data set 060721cpa1.*

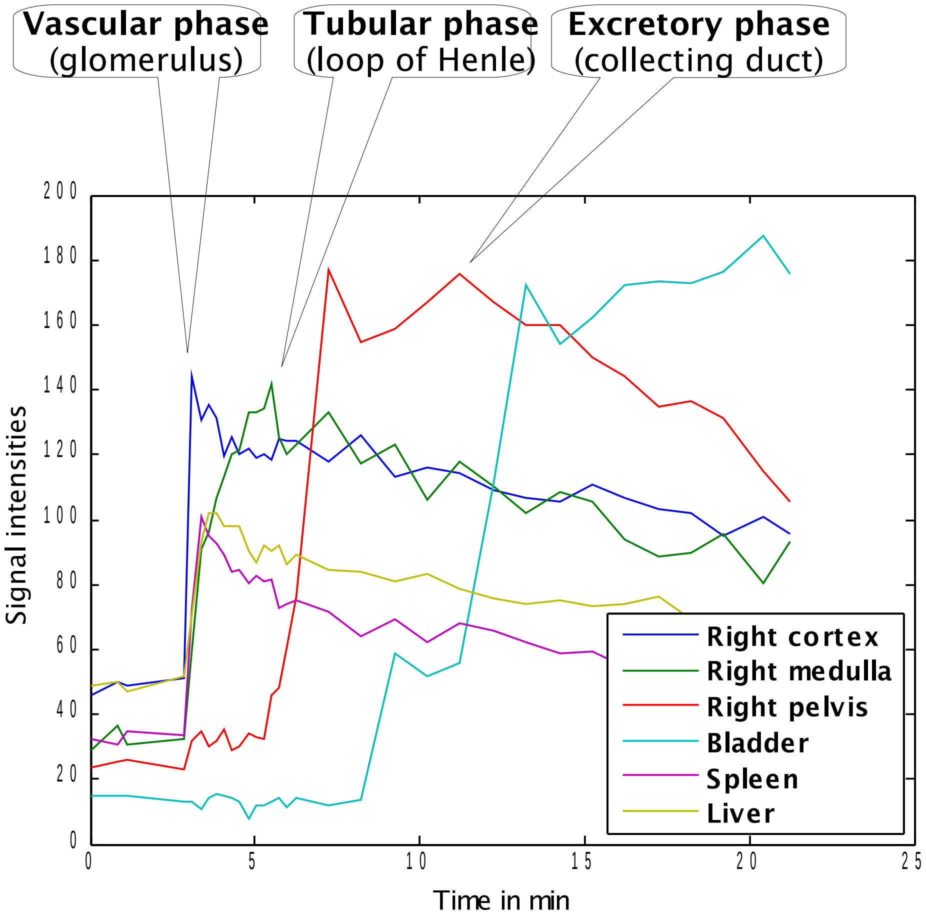


Figure 14: Pattern of signal intensities

- *Voxels representative of the different physiological tissue types are selected from Figure 13. Their signal intensities are averaged with those of their 8 neighbours and plotted over time.*
- ***Pre-enhancement phase** (t=2.9min, see also Figure 13) shows low signal intensities before contrast agent uptake except usually for liver which has a short T1 relaxation time.*
- ***Vascular phase** (t=3.1min, see also Figure 13):*
  - *Renal cortex enhances at the outer rim and reaches a peak in intensity.*
  - *Spleen enhances heterogeneously*
  - *Liver will enhance slightly later due to the delay required to deliver Gd-DTPA from the portal venous system.*
- ***Tubular phase** (t=5.5min, see also Figure 13):*
  - *Renal medulla is first hyperintense and reaches a signal intensity peak*
  - *Then, the renal parenchyma (both cortex and medulla) becomes isointense.*
- ***Excretory phase** (t=13.2min, see also Figure 13) shows increasing signal intensity in renal pelvis and bladder.*
- *Left kidney shows delayed excretion revealing a possible obstruction (see Figure 13).*
- *Data set 060721cpa1, slice 6.*

- A **tubular phase** where the Gd-DTPA moves down to the medulla via the descending limb of the loop of Henle and gets back to the cortex via the ascending part and the DCT.
    - The first part of this phase is identified by a contrast inversion: the signal intensity of the renal cortex decreases due to glomerular filtration and dilution in the PCT whereas that of the renal medulla increases as long as the inflow of Gd-DTPA from the cortical vasculature into the tubules exceeds the outflow into the CD. The medulla is then hyperintense and reaches a peak in intensity which is slightly higher and occurs later than that of the cortex, reflecting the concentration of Gd-DTPA along the tubules and a maximal concentration in the LOH.
    - Subsequently, the second part of the tubular phase is characterised by a decline but an equilibrium of Gd-DTPA flow between the kidney cortex and medulla. Their signal intensities both decrease but become isointense resulting in a homogeneously enhanced parenchyma over time. This can be explained by the flow of Gd-DTPA within the ascending limb of the loop of Henle and the DCT which are impermeable to water and keep the concentration of contrast agent approximately constant.
    - The contrast agent leaves the spleen and the liver respectively via the splenic vein and the hepatic vein and thus, undergo a sharp decrease of their signal intensities.
- An **excretory phase** where the signal intensities of renal cortex and medulla both keep decreasing homogeneously due to the contrast elimination into the CD. It results in a brighter signal intensity first in the kidney pelvis, which is the most intense region during a normal study because of urine accumulation. Then, contractions occur 2 to 6 times per minute to form and transport, via the ureter, bolus of urine into the bladder which becomes also brighter.

Finally, it is interesting to notice a few other points:

- The vascular phase is the one that most of the previous studies has focused on to segment kidneys because it offers a good cortico-medullary differentiation with hypointense signals for other abdominal organs. But this difference fades away in the subsequent phases.

- Compared to the spleen and the liver, the kidney enhances the most intensively. This can be explained by its prominent vascularity and its ability to filter and concentrate Gd-DTPA.
- On T1-weighted images, the liver signal intensity displays more signal intensity in the pre-enhancement phase compared to other abdominal tissues due to its characteristic short T1 relaxation time (Figure 15) ([53]).
- On T1-weighted images, the signal intensity of the spleen is usually lower than that of the liver [20] but is higher in terms of relative contrast enhancement, which reflects the prominent vascularity of the spleen [12] and the short T1 relaxation time of the liver (that is, an additional decrease in T1 relaxation time caused by a contrast agent leads only to a slight increase of liver signal intensity compared to other tissues).
- The renal enhancement pattern was explained by the T1 shortening effects of Gd-DTPA uniquely (that is, an increased concentration of Gd-DTPA yields an increased signal intensity) but it must be kept in mind that T2 shortening effects cannot be neglected and often occur in the renal medulla and pelvis.
- Alterations in the normal enhancement pattern can be observed in cases of uropathies like obstruction.

Tissue	T1 relaxation time (msec) at 1.5 Tesla
Kidney	
Cortex	966
Medulla	1412
Liver	586
Spleen	1057

Figure 15: Average T1 relaxation times at 1.5 Tesla [53]

Thus, the kinetics of contrast uptake enhance differently depending on the abdominal tissues and they describe a reproducible enhancement pattern which can be explained by the anatomical and physiological differences between tissues. Consequently, this can be a strong feature to segment the kidneys from the surrounding tissues as well as to segment the different renal compartments (cortex, medulla and pelvis).

2.4 Technical issues

2.4.1 Quantification of contrast agent

The paramagnetic effects of Gd-DTPA (or any other paramagnetic contrast agent) on relaxation times T1 and T2* (or T2) of water protons can be described by the following equations:

$$\begin{aligned} \frac{1}{\mathrm{T1}} &= \frac{1}{\mathrm{T1}_0} + r_1 \cdot [\mathrm{Gd}] \\ \frac{1}{\mathrm{T2}^*} &= \frac{1}{\mathrm{T2}_0^*} + r_2 \cdot [\mathrm{Gd}] \end{aligned} \tag{2.4.1}$$

- $\mathrm{T1}_0$ and $\mathrm{T2}_0^*$ are the tissue relaxation times in the absence of Gd-DTPA.
- r_1 and r_2 are relaxivity constants specific to Gd-DTPA.
- $[\mathrm{Gd}]$ is the concentration of Gd-DTPA.

Moreover, the relationship between the relaxation times and the MR signal S_{MR} on spoiled gradient echo images can be described by the following equation [34]:

$$S_{MR} = M_0 \cdot \sin\alpha \cdot \mathrm{e}^{-\frac{\mathrm{TE}}{\mathrm{T2}^*}} \cdot \frac{1 - \mathrm{e}^{-\frac{\mathrm{TR}}{\mathrm{T1}}}}{1 - \cos\alpha \cdot \mathrm{e}^{-\frac{\mathrm{TR}}{\mathrm{T1}}}} \tag{2.4.2}$$

- M_0 is the equilibrium longitudinal magnetization.
- α is the flip angle.
- TE and TR are respectively the echo time and the repetition time.

Equations (2.4.1) shows that an increase of Gd-DTPA concentration shortens both the relaxation times T1 and T2* and hence modify the MR signal through equation (2.4.2).

It is interesting to notice that if T2* is long relatively to TE, then the T2* shortening effects can be neglected in equation (2.4.2) and the MR signal can be approximated by:

$$S_{MR} = M_0 \cdot \sin\alpha \cdot \frac{1 - \mathrm{e}^{-\frac{\mathrm{TR}}{\mathrm{T1}}}}{1 - \cos\alpha \cdot \mathrm{e}^{-\frac{\mathrm{TR}}{\mathrm{T1}}}} \tag{2.4.3}$$

Assuming a given spoiled gradient echo sequence with α=15°, TR=6.2 msec and TE=2.8 msec, Figure 16 simulates the complex relationship between the MR signal and the Gd-DTPA concentration in medullary tissues by using equations (2.4.1), (2.4.2), (2.4.3) and values reported in the literature ($T1_0$=1412 msec, $T2_0$*=85 msec [53], r_1=4.5 kg mmol^{-1} sec^{-1}, r_2=30 kg mmol^{-1} sec^{-1} [18] [37]).

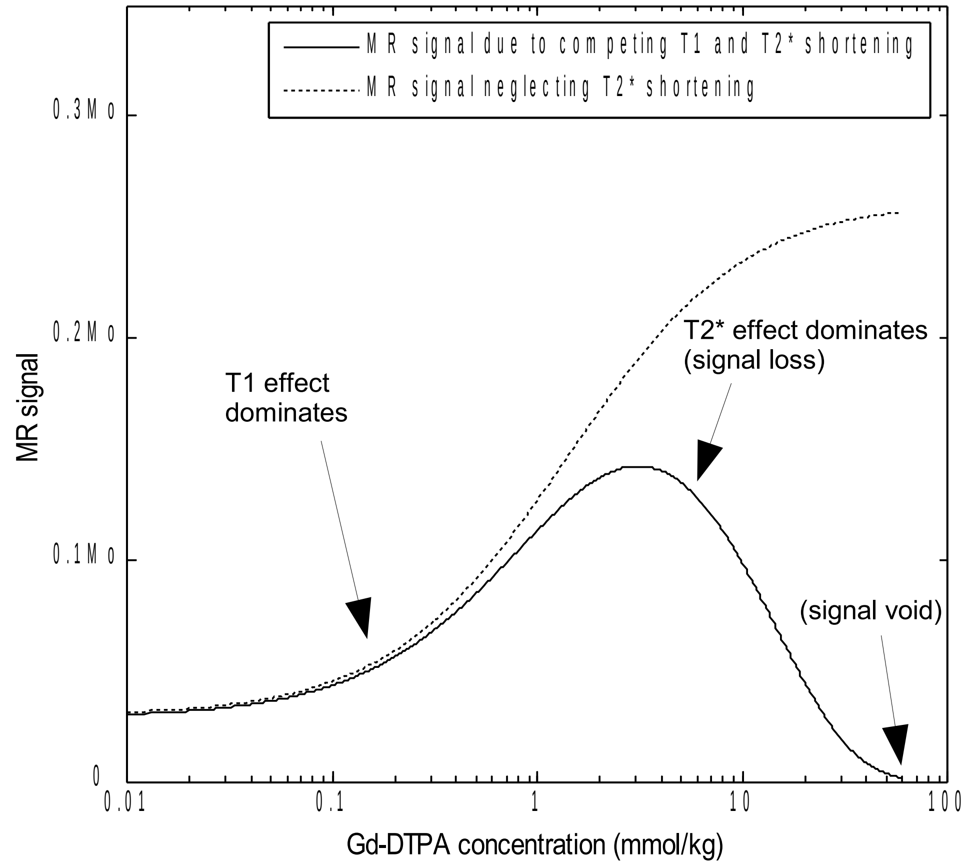


Figure 16: MR signal as a function of Gd-DTPA concentration in medullary tissues

It can be concluded from the simulation that the MR signal is not linearly related to the concentration of Gd-DTPA present in the tissue. Indeed, the MR signal is determined by the competing paramagnetic effects of T1 and T2* shortening by Gd-DTPA (Figure 16):

- At low Gd-DTPA concentrations, T1 shortening effect dominates and enhances the signal intensity of urine.
- But at high concentrations, T2* shortening effect dominates, resulting in a decreased rather than increased signal intensity of urine.

T2* shortening effect is clearly undesirable because it decreases the positive enhancement of the urine and leads to a signal loss or eventually to a signal void. In the simulation, the T2* effect alters significantly the MR signal when the concentration of Gd-DTPA rises locally above 1 mmol/kg and causes a signal void above 60 mmol/kg. This often occurs in DCE-MRU because renal medulla and pelvis can locally concentrate and accumulate Gd-DTPA up to 100-fold. Consequently, the concentration of contrast agent in DCE-MRU cannot be simply and uniquely determined on the basis of signal intensities.

This is in contrast to nuclear scintigraphy where the signal intensity measures the count of gamma photons in the tissue of interest and hence, is directly proportional to the concentration of radionuclides.

To overcome the issue in DCE-MRU, a few solutions have been proposed:

- A lower concentration of Gd-DTPA can be used to prevent gadolinium from being too much concentrated in the renal medulla and collecting system and thus, avoid the dominant T2* effect. Studies ([23]) suggest that doses lower than 0.1 mmol/kg may keep the relationship between signal intensity and Gd-DTPA concentration approximately linear.
- Furosemide can be administered prior to Gd-DTPA. It inhibits water reabsorption in the loops of Henle and hence, dilutes the concentration of Gd-DTPA. It also causes an uniform distribution of Gd-DTPA inside the urinary tract [25].
- The parameters of the spin echo sequence and gradient echo sequence could also be adjusted to the concentration of Gd-DTPA that is expected to be found in the kidneys so that T2* effects can be reduced ([23] [51]).

2.4.2 Motion artefacts

Patient motions during image acquisition must be avoided because they alter the signal and degrade the image quality, especially the cortico-medullary differentiation.

- Gross patient motion causes blurring of the image. This happens when the patient does not keep still, for example, in response to the contrast agent administration. Or the patient may feel some discomfort, especially because a MR scanner is very noisy and the acquisition can take a long time (usually more than 20 min with our imaging protocol).
- Physiological motion like respiration can produce blurring and ghost images. To avoid respiratory artefacts, the image acquisition can be done repetitively during several breath holds. But even in this case, the kidneys may slightly shift spatially or rotate from one image volume to another if the diaphragm is not held in the same position for each acquisition. Kidney motion is a major problem if the perfusion sequence needs to be tracked over time for renogram or segmentation methods like ours. For non cooperative patients, especially children under 10 years old, sedation is required during the image acquisition.
- Motion of blood containing contrast agent causes ghost artefacts which are equally separated (Figure 17).
- Differential mixing of concentrated and less concentrated contrast and unopacified urine in the renal pelvis in the early phase of contrast excretion from the renal medulla causes streaks that appear to come out of the renal pelvis (Figure 18).

It is especially important to avoid kidney motion in DCE-MRU. Indeed, all clinical MRU examinations require an accurate positioning of distinct VOIs over the different renal compartments to generate a reliable renogram; but it cannot be simply done by just copying the VOIs over all the image volumes of the sequence because kidneys may slide or rotate from one image volume to another due to unavoidable patient motion during the image acquisition (Figure 19). Therefore, a registration procedure is often a requirement to correct for complicated 3-D motion of the kidneys and align them over the sequence.

Registration is even more necessary for our segmentation techniques: if the data sets were not corrected against kidney motion, it would not be possible to track correctly the renal signal intensity over the perfusion sequence and thus, our segmentation techniques, based on the differences in kinetics of contrast uptakes between different physiological tissues, would fail. So, the accuracy of our kidney segmentation techniques will strongly depend on the quality of the prior registration.

Registration is not the purpose of this project and thus, our segmentation techniques will only be tested on data sets which could reasonably be corrected against translation via a method based on cross correlation.

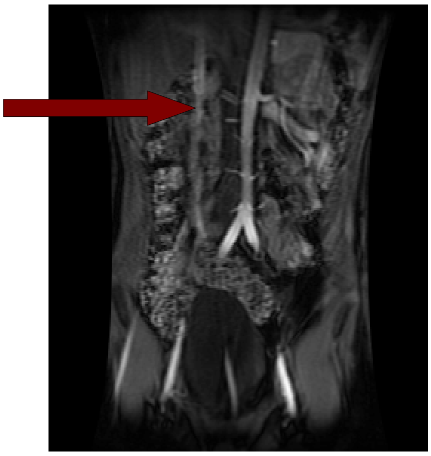


Figure 17: Ghost image of the aorta

- *The ghost image is arrowed.*
- *Data set 040902pa1, slice 14, t=3.8 min.*

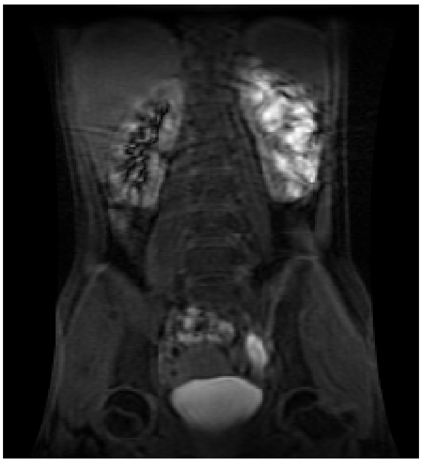
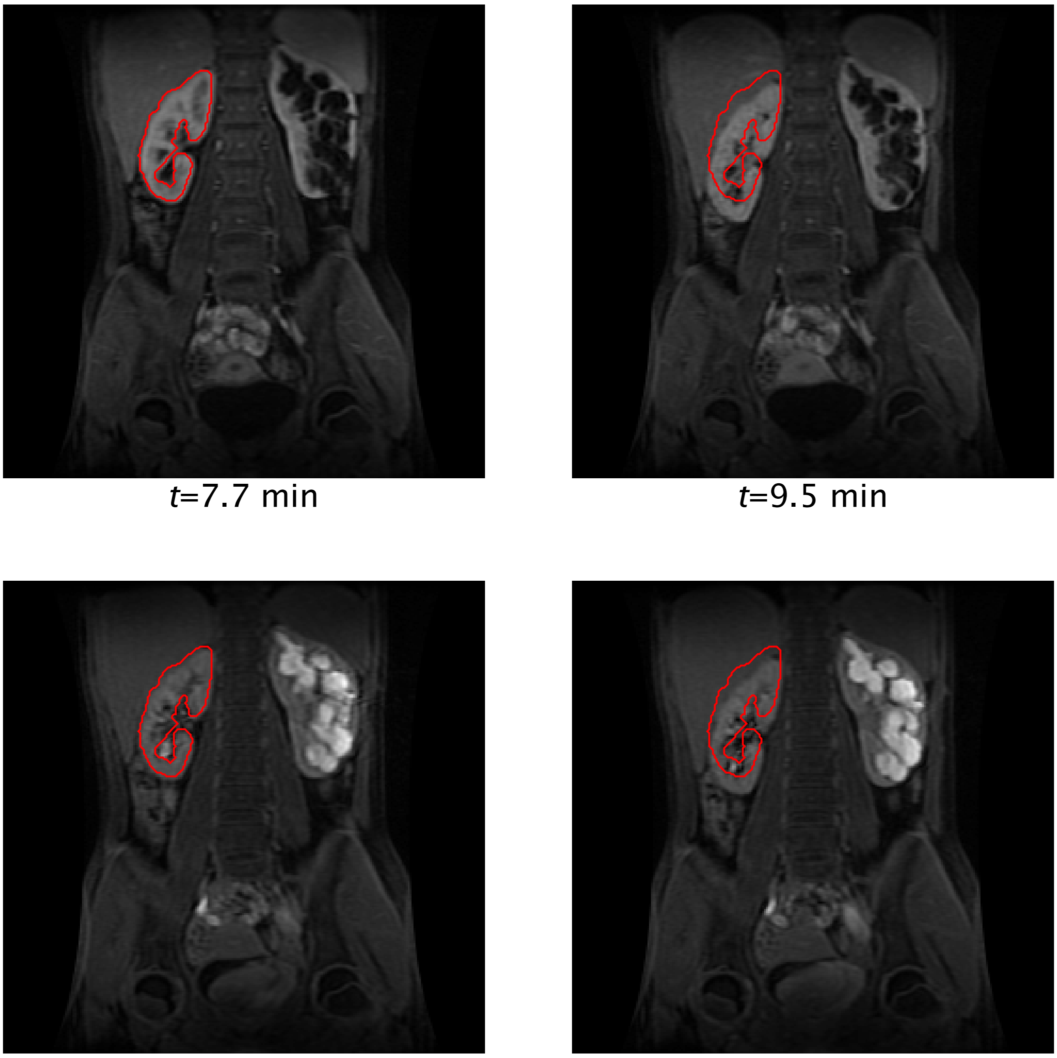


Figure 18: Visible streaks appear to come out of both renal pelvis

- *Data set 040831cpa1, slice 9, t=29.4 min.*



t=7.7 min

t=9.5 min

t=23.4 min

t=24.4 min

Figure 19: Kidney motion in DCE-MRU

- *A region of interest is placed manually over the right kidney, on slice 8, at t=7.7min.*
- *The same region of interest is copied on the same slice 8 over the perfusion sequence (t=9.5min, t=23.4min and t=24.4min).*
- *The sequence was acquired during multiple breath holds. But clearly, because the right kidney shifts and deforms between acquisitions, unwanted areas of surrounding tissues like liver are included and will contribute to the region of interest if a renogram is generated.*
- *Data set 040831cpa1, slice 8.*

2.4.3 Partial volume artefacts

Partial volume artefacts occur for voxels which contain a mixture of multiple tissue types. The mixed voxels have then a signal that is a weighted sum of the signals from the multiple tissue types. It happens especially at tissue boundaries if the spatial resolution is insufficient (even though a slice contains 256×256=65536 voxels), resulting in blurring and ambiguities in those regions.

Segmentation algorithms can take into account those uncertainties by allowing different tissue types to overlap (see our proposed segmentation methods in section **4.**).

2.5 Scanning protocol

All the 20 patients were children. A 3-D Fast Spoiled Gradient Recalled sequence with a flip angle $\alpha=15°$, TR=6.2 msec and TE=2.8 msec was performed. The reconstructed slice thickness was approximately 5 mm and the field of view (FOV) was 300mm×150mm. The acquisition matrix was 256×128 giving 300/256, that is approximately 1.2mm×1.2mm×5mm voxels.

The 3-D acquisition scans were repeated at an interval of approximately 16 sec for the first 5-6 min and subsequently, they were repeated at 1 min intervals for about 20 min. Several pre-contrast acquisitions were usually performed and the amount of Gd-DTPA that was intravenously injected was 0.1 mmol/kg of subject weight (volume of 2mL/10kg).

The gradient echo sequence is favoured over the spin echo sequence because of its shorter scanning times. Thus, artefacts due to patient motions can be reduced and 3-D acquisitions are made more practical. Moreover, 3-D acquisitions can be obtained with a single breath hold and there is thinner slice thickness with no interslice gap.

Spatially varying T2*-weighting is reduced by employing a spoiling process to eliminate after each echo the residual transverse magnetization.

Faster scans and better temporal resolution would be possible if a single slice only was acquired. The whole 3-D volume of the abdomen was wanted. Moreover, nowadays, with current technology, a 3-D volume could be acquired in as little as 4-5 sec.

A brief description of our 20 patient MR data sets are given in **Appendix**.

3 Related works in kidney segmentation

Previous kidney segmentation techniques can be classified in two groups, depending on whether they are based on spatial information or temporal information:

- In the spatial approach, segmentations are performed on one image volume chosen at one time point, usually the one when the cortico-medullary differentiation is optimal, that is during the vascular phase.
- In the temporal approach, segmentations differentiate tissues according to their contrast agent uptake over time by considering the intensity time course associated to the voxels.

3.1 Methods based on spatial information

3.1.1 Methods based on deformable models

Tsagaan *et al.* [48] introduces a priori knowledge of the kidney shape to segment 3-D volumes of the kidneys in 33 CT images: information as average and statistical variation of the kidney's shape is derived from a training set and incorporated as an energy function into a deformable model represented by Non-Uniform Rational B-Spines (NURBS) surface. The size and the location of the model in the CT images is initialized automatically, then the model is deformed to match the kidney target by minimizing the energy function and finally, the renal hilum is extracted in a post-processing stage by region-growing.

Rao *et al.* [60] exploits a similar idea using deformable shape models called m-reps to automatically segment 3-D kidneys in 12 CT images. M-rep models are obtained by using a set of about 50 training images where kidneys are manually segmented. After the kidney m-rep model is first interactively initialized over the target kidney in the CT images, an automatic and iterative procedure, driven by an objective function, positions it closer to the target via a global similarity transformation and changes its shape to match the target based on principal geodesic analysis until no changes are observed.

However, deformable model approaches that use a priori information to match the model against a target object usually achieve sophisticated segmentation results provided that the target object has a predictable shape. This is the case for normal anatomic structures which can thus be represented by a typical shape and serves as a model (for example, the m-rep model in [60] is merely a mean shape). But it is believed that deformable models are not the right approach for kidney segmentation in MRU because the target organ can have very various shapes, especially if it is malfunctioning, and thus reliable shape information cannot be incorporated. Moreover, the authors in [48] also note that their segmentation method fails if the target kidney shape is too deformed by surrounding organs such as the liver or the spleen.

3.1.2 Methods based on the use of contrast agent

Nuclear scintigraphy

Tomaru *et al.* [27] is a semi-automated method based on two thresholds to assign a renal region of interest (ROI) in ^{99m}Tc-MAG3 (Technetium Mercaptoacetyltriglycine) renography. The operator is required to click on the centre of the kidney. Then, the maximal renal count inside a small rectangular area centred on the clicked pixel is calculated. The pixels, inside a larger area centred again on the clicked pixel, whose count is greater than 60% of the maximal count are highlighted to form areas. The only area that includes the clicked pixel is disconnected from the others and is called the renal central area. Finally, pixels that are adjacent to the renal central area and whose count is greater than 30% of the maximal renal count are chosen to form the renal peripheral area. The renal central area and the peripheral area constitute the renal ROI. The authors show that their method detects renal contour well and calculates reproducible renal uptake of ^{99m}Tc-MAG3.

Dynamic Contrast Enhanced-MRI

Several approaches in DCE-MRI take the advantage of the perfusion sequence by segmenting the kidneys at a time point when they are especially enhanced by the contrast agent.

De Priester *et al.* [39] proposes a semi-automated method on dynamic MR image series of renal transplant to segment the kidney on 2-D slices and obtain MR renographs of the cortex and medulla by placing two ROIs that overlay the peripheral and central renal parenchyma. The process first involves subtraction of the precontrast images from the early-enhanced images, thresholding to yield a binary image in which there are only pixels whose signals enhance, removing of small objects to keep a single area that corresponds to the renal parenchyma, morphological closing and erosion to obtain both the renal parenchyma and the hilar structures. Then, two bands are automatically generated: a first peripheral band overlaying the renal cortex and hilum and a second central band composed of cortical tissue, medullary tissue and hilar structures. Finally, an operator delineates the non-hilar area to define, from the two previous bands, a peripheral ROI and a central ROI which are copied to all the time points of the dynamic series. Cortical signal curves can be obtained from the peripheral ROI while medullary signal curves can be computed by a weighted subtraction of the mixed cortico-medullary signal of the central ROI from the cortical signal. However, this method still needs further studies to extract 3-D volumetric native kidneys.

In a study by Coulam *et al.* [43], an operator first traces manually the kidney perimeter on a few two-dimensional DCE-MR slices acquired during cortical enhancement phase and nephrographic enhancement phase. Subsequently, an interpolation between the traced slices is computed to obtain automatically a cortical phase volume and a nephrographic phase volume of the kidneys. Finally, a lower signal intensity threshold is selected interactively to isolate, on the one hand, the renal cortex from the cortical phase volume and, on the other hand, the parenchyma from the nephrographic phase volume. The medulla could also be segmented by just subtracting the cortex from the parenchyma. This approach was performed on eight anaesthetized pigs.

Koh *et al.* [65] present a method that uses the morphological 3-D h-maxima transform to segment kidneys on 5 DCE-MRI studies. Rectangular masks are first generated to locate the kidney regions wherein a Canny edge detector is subsequently performed over the slices to capture the shape of the kidneys. The result is incorporated into the output of the 3-D h-maxima transform as a spatial constraint to minimize leakage. The authors show that their approach extracts the kidney cortex without the unwanted areas of the surrounding tissues. However, their method depends heavily on the generated constraint: there are cases when a kidney is mistaken by a surrounding organ like the spleen for example or when the Canny edge detector does not perform well. Finally, few remarks must be made about their work:

- Their proposed segmentation method was processed on a limited selection of studies where the 3-D h-maxima transform, assisted by manual cropping as required, could yield relatively good results.
- 3-D h-maxima transform alone can work reliably with little operator intervention in kidneys of normal function.
- 3-D h-maxima transform and the Canny edge detector are not sufficiently robust in the data sets which have been selected for this work.

Unfortunately, all those previous described methods need the operator interaction ([39] [43]) or are dependent on parameters whose values can be quite subjective (appropriate threshold values are needed for [27] [39] [43], and appropriate values for the edge detector and the 3-D h-maxima transform are needed for [65]). Furthermore, they all give limited precision in differentiating structures (like right kidney from liver for [27], cortex from medulla for [43]) if the contrast is insufficient or if the signal-to-noise ratio is poor. Especially, a poorly functioning kidney often yields a suboptimal contrast between renal cortex and medulla even during the vascular phase.

3.2 Methods based on temporal information

While many existing spatial segmentation approaches can be applied, the use of temporal information is, to our knowledge, still limited.

Boykov *et al.* [38] proposes a semi-automated method that introduces a temporal Markov Model into a graph cut algorithm to segment dynamic MRI volumes. The nodes of the graph correspond to the voxels and the edges correspond to the relationship between the neighbouring voxels. The aim is to find the cut in the graph that generates an optimal segmentation imposed by the edge weight values determined from hard and soft constraints:

- Hard constraints are provided by an operator who marks certain voxels that absolutely have to be part of the volume of interest and certain pixels that have to be part of the background.
- Soft constraints are defined as a cost function that incorporates both boundary and region information.

Then, the algorithm computes, via a min-cut/max-flow algorithm, the cut in the graph which corresponds to the global minimum of the cost function among all the possible segmentations that satisfy the hard constraints imposed by the user. That is, the cut gives the extraction of the volume of interest from the background. Their method is applied to segment a registered perfusion sequence of one single kidney volume in three iterations: the entire kidney is first extracted from the background, then the medulla is separated from the cortex and the collecting system, and finally, the collecting system is cut from the cortex.

Sun *et al.* [47] minimizes an energy function that takes into account both the spatial correlation of the pixels in the same image and their temporal correlation through the image sequence to segment in 2-D kidney cortex of 10 normal and 10 transplanted rats in MRI sequences enhanced by USPIO (Ultrasmall Superparamagnetic Iron Oxide). It is assumed that the image sequence is formed of an inner region and an outer region whose dynamic profiles are different from each other. The energy function measures the dissimilarity between the dynamic signals and the average dynamic signal in each of those two regions and is minimized by a level set-based approach or a region growing-based approach.

3.3 Summary

The flaw of all intensity-based segmentation methods is their over-sensitivity to noise and artefacts. They are best suited for normally functioning tissues or images where the voxel intensity and edge information are not degraded.

The evaluation of the signal intensities provides useful physiological information that can help achieve better differentiation of organs in MRU. But previous methods based on temporal information still need the intervention of an operator (hard constraints from the operator is required for [38]) or do not consider all three spatial dimensions ([47]) to extract kidney volumes. Furthermore, they still need a clinical validation.

4 Proposed methods for kidney segmentation

4.1 Definitions and notations

4.1.1 Data set

A volumetric data set of N_s slices of $N_i \cdot N_j$ pixels is represented in a vector form of $\mathbb{N} = N_i \cdot N_j \cdot N_s$ components called voxels.

The dynamic volumetric data set $\mathbb{D}$ of T time points, referenced as four dimensional *(t* + 3-D or 4-D) data set, is represented in a matrix form of $\mathbb{N}$ rows and T columns. Each row of the matrix $\mathbb{D}$ represents the time intensity curve (TIC) of a given voxel and each column represents the intensity of all the voxels at a given time point.

A TIC associated to a voxel is also called a quadrixel and the intensity of all the voxels at a given time point forms a volume image. The 4-D dataset $\mathbb{D}$ can then be entirely defined by the $\mathbb{N}$ quadrixels $\boldsymbol{d}_{n.}$ of T time points or by T volume images of $\mathbb{N}$ voxels:

$$\mathbb{D} = \begin{bmatrix} d_{11} & \cdots & d_{1T} \\ \vdots & \ddots & \vdots \\ d_{\mathbb{N}1} & \cdots & d_{\mathbb{N}T} \end{bmatrix} = \begin{bmatrix} \boldsymbol{d}_{1.}' \\ \vdots \\ \boldsymbol{d}_{\mathbb{N}.}' \end{bmatrix} = \begin{bmatrix} \boldsymbol{d}_{.1} & \cdots & \boldsymbol{d}_{.T} \end{bmatrix} \tag{4.1.1}$$

$$\forall n = 1 \ldots \mathbb{N},\ \boldsymbol{d}_{n.} = [d_{n1} \quad d_{n2} \quad \ldots \quad d_{nT}]' \text{ is the } n^{th} \text{ quadrixel} \tag{4.1.2}$$

$$\forall t = 1 \ldots T,\ \boldsymbol{d}_{.t} = \begin{bmatrix} d_{1t} \\ d_{2t} \\ \ldots \\ d_{\mathbb{N}t} \end{bmatrix} \text{ is the volume image at time point } t \tag{4.1.3}$$

Similarly, a trixel can also be defined as the TIC associated to a pixel for a dynamic planar data set referenced as a three dimensional data set *(t* + 2-D*).*

The dynamic volumetric data set $\mathbb{D}$ is supposed to be perfectly registered over time.

4.1.2 Time enhancement curve (TEC)

While the time-intensity curve (TIC) is the direct measured raw signal intensities, the time-enhancement curve (TEC) is defined and can be regarded as a normalized TIC or a relative signal enhancement. For a given voxel n , it is expressed as:

$$\boldsymbol{e_n} = [e_{n1} \quad e_{n2} \quad \dots \quad e_{nT}] \tag{4.1.4}$$

- □ $e_{nt} = \frac{d_{nt} - \bar{d}_{n\tau}}{\bar{d}_{n\tau}}$ is the TEC value at time point t , with $\bar{d}_{n\tau} \neq 0$.
- □ $\bar{d}_{n\tau} = \frac{1}{\tau} \sum_{t=1}^{t=\tau} d_{nt}$ is the average pre-contrast TIC defined by the first τ time points.

TIC retains information about the amplitude of the measured raw signal intensities while TEC emphasizes the relative enhancement of the signal.

All our proposed methods based on clustering analysis and factor analysis consider the TIC associated to the voxels, also called quadrixels. However, the use of TEC will later be justified in clustering analysis to remove the background composed of non enhanced voxels.

4.2 Preliminary concepts for our approaches

4.2.1 Cluster analysis

Motivations for a cluster analysis approach

In a cluster analysis approach, a given quadrixel $\boldsymbol{d}_{n.}$ of T components is viewed as one data point in an Euclidean space of T dimensions that shall be called data space or feature space.

If all the N quadrixels of a study D were mapped into the data space of T dimensions, high density areas of data points, called clusters, would be found, reflecting different groups of quadrixels having similar TIC.

Thus, kidney quadrixels, which share the same temporal characteristics of contrast agent uptake, would plot inside a certain cluster c_{k^*} while other tissue quadrixels would plot inside other different clusters c_k, $k \neq k^*$. The aim of this segmentation approach can then be considered as a partition of all the quadrixels into K different physiological regions represented by K clusters when mapped into a T-dimensional space. The clustering process is based on a criterion of similarities between quadrixels: data points in a same cluster c_{k^*} must be more similar to each other than to those in c_k, $k \neq k^*$.

K-Means Clustering

For illustration purposes, the most commonly used clustering algorithm, called K-means clustering [35], is considered to find in a 4-D data set D, K physiological regions of similar quadrixels $\boldsymbol{d}_{n.}$.

K-means clustering is an iterative process that minimizes the sum-of-squared error criterion:

$$E(\boldsymbol{W})=\sum_{k=1}^{K} \sum_{\boldsymbol{d}_n \in c_k} \|\boldsymbol{d}_{n.}-\boldsymbol{w}_k\|^2 \qquad (4.2.1)$$

- K is the number of clusters and represents the number of different physiological regions that must be segmented.

- c_k is the k^{th} cluster and represents the segmentation of a given physiological region.
- $\boldsymbol{w}_k$ is the centre of cluster c_k and is sometimes called prototypes. It can be calculated by averaging all the TIC that belong to region c_k.
- $\boldsymbol{W}=[\boldsymbol{w}_1 \ \ldots \ \boldsymbol{w}_K]$ are the K cluster centres or prototypes.

From (4.2.1), the centres $\boldsymbol{w}_k$ of clusters c_k, $k=1\ldots K$, can be interpreted as being the best representation of the data set $\mathcal{D}$ in the sense that it minimizes:

$$\sum_{\boldsymbol{d}_n \in c_k} \|\boldsymbol{d}_{n.} - \boldsymbol{w}_k\|^2 \tag{4.2.2}$$

Hence, (4.2.1) measures the total square error in representing the given data set $\mathcal{D}$ by the K cluster centres $\boldsymbol{W}=[\boldsymbol{w}_1 \ \ldots \ \boldsymbol{w}_K]$.

That is, the cluster centres $\boldsymbol{W}=[\boldsymbol{w}_1 \ \ldots \ \boldsymbol{w}_K]$ are the typical and best representatives TIC of each of the corresponding K physiological regions.

However, K-means clustering has limitations that are well-known. It has a strong tendency to form clusters of roughly equal size and is best suited for data sets that are essentially compact, hyper-spherical in shape and well-separated from each other. But this is often not the case for our 4-D MRU studies because clusters are usually blurred and close to each other due to noise, different artefacts and partial overlaps that affect TIC.

Proposed Clustering methods

Three other clustering methods for kidney segmentation in 4-D MR images are proposed:

- **Kohonen Neural Network** (KNN) which is also a hard clustering method in the sense that it assigns each data point to exactly one single cluster (**Chapter 4.4**).
- **Fuzzy K-means Clustering** (FKM) which is a soft clustering version of K-means clustering in the sense that it incorporates the concept of fuzzy membership functions of a data point to multiple clusters (**Chapter 4.5**).

- **Expectation-Maximization Clustering** (EM) which can also be considered as a soft clustering as it is a statistical approach that assigns each data point to multiple clusters with a certain probability (**Chapter 4.6**).

4.2.2 Factor analysis

Factor analysis is a very different kind of approach. It aims at extracting the relevant signal from the noise and then summarizing it into a set of a reduced number of variables called factors.

Applied to medical image sequences, it incorporates further constraints to yield factors that can describe physiologically the dynamic and the spatial distribution of independent structures. Indeed, the proposed factor analysis of medical image series (FAMIS) expresses TIC as a linear decomposition of a limited number of fundamental factor curves and associated factor volumes which represent respectively the temporal behaviour and the spatial distribution of the contrast agent in different regions of the abdomen.

The details of the proposed FAMIS with new constraints and similar works by other groups will be explained in **Chapter 4.7**.

4.3 Preprocessing: background elimination

This stage consists of background elimination. There are two types of background which can be considered for elimination:

- The background composed of voxels that do not enhance over the image acquisition. Those voxels belong to tissues where the contrast agent uptake does not occur and thus are unlikely to be renal voxels. To eliminate this background, two strategies are possible:
    - A threshold on the TEC: each quadrixel $\boldsymbol{d}_{n.}$ of the original data set D, $n=1\ldots\mathsf{N}$, whose mean TEC value over time, defined as $\frac{1}{T}\sum_{t=1}^{T} e_{nt}$, is below a selected enhancement threshold $\Theta_E \geqslant 0$, is excluded. This is the strategy chosen for our clustering analysis methods.
    - A threshold on the TIC: each quadrixel $\boldsymbol{d}_{n.}$ of the original data set D, $n=1\ldots\mathsf{N}$, whose mean TIC value over time, defined as $\frac{1}{T}\sum_{t=1}^{T} d_{nt}$, is below a selected intensity threshold $\Theta_I \geqslant 0$, is excluded. This is the strategy chosen for our factor analysis method.
- There is also a background composed of voxels that do enhance but are not of interest because they do not belong to renal tissues. This is especially the case for the abdominal aorta and the iliac arteries for example. Although the renal cortex enhancement phase persists beyond the initial arterial peak in the large arteries, this difference may not be enhanced enough with the temporal resolution used, and as a consequence, the cortex and some arteries, like the abdominal aorta and the iliac arteries, may exhibit TIC which will not be differentiated by the proposed segmentation approaches. The strategy that is adopted requires an operator to manually draw a region of interest (ROI) on a maximum intensity projection (MIP) derived from the dynamic sequence to exclude those tissues that may be problematic. Among our 20 MR data sets, this option is used only once.

Let $\boldsymbol{D}=\{\boldsymbol{d}_{1.}, \boldsymbol{d}_{2.}, \ldots, \boldsymbol{d}_{N.}\}$ be the data set whose background is removed by the preprocessing stage and which is now comprising of only $N \leqslant \mathsf{N}$ quadrixels that are submitted for further segmentation analysis.

4.4 Kohonen Neural Network (KNN)

4.4.1 Introduction

A Kohonen Network is a self-organizing network of artificial neurons that learns the characteristics of the input data points in a competitive way and thus, is often known as a competitive network. Each output neuron stores a weight vector that can be viewed as the centre of a cluster in the input data space.

Wismuller *et al.* [67] extended the Kohonen Neural Network to a Self-Organizing Map to discover clusters of similar temporal behaviours on dynamic cerebral contrast-enhanced perfusion MRI time-series.

4.4.2 Network architecture

It is a single-layer network where the inputs are fully connected to the output (competitive) layer (Figure 20). The network is built with T inputs and K output neurons. To recall, T is the number of time points and K is the number of clusters.

The centre of cluster c_k is represented by a prototype $\boldsymbol{w}_k = [w_{k1} \quad w_{k2} \quad \dots \quad w_{kT}]'$ that is attached to each of the K output neurons. The K prototypes $\boldsymbol{W} = [\boldsymbol{w}_1 \quad \boldsymbol{w}_2 \quad \dots \quad \boldsymbol{w}_K]$ can be also be interpreted as the weights between the T inputs and the K output neurons.

4.4.3 Training

Description

The training refers to finding the cluster centres $\boldsymbol{W} = [\boldsymbol{w}_1 \quad \boldsymbol{w}_2 \quad \dots \quad \boldsymbol{w}_K]$ such that the data set $\boldsymbol{D} = \{\boldsymbol{d}_{1.}, \boldsymbol{d}_{2.}, \dots, \boldsymbol{d}_{N.}\}$ is labelled into K groups representing physiological regions of similar temporal behaviours.

During training, an input quadrixel $\boldsymbol{d}_{n.} = [d_{n1} \quad d_{n2} \quad \dots \quad d_{nT}]'$ is chosen randomly and presented one at a time to the network.

At each presentation, the euclidean distances h_k between the input quadrixel and each of the K weight vectors are computed:

$$h_k = \left\| \boldsymbol{w}_k - \boldsymbol{d}_{n.} \right\| = \sqrt{(\boldsymbol{w}_k - \boldsymbol{d}_{n.})'(\boldsymbol{w}_k - \boldsymbol{d}_{n.})} \text{ with } k = 1 \ldots K \tag{4.4.1}$$

The output neuron k^* whose weight vector $\boldsymbol{w}_{k^*}$ is nearest to the input quadrixel $\boldsymbol{d}_{n.}$ is considered as the winning neuron and outputs $u_{k^*}=1$ while others output $u_k=0$, $k \neq k^*$:

$$u_{k^*} = \begin{cases} 1 \text{ if } h_{k^*} = \underset{k}{\operatorname{argmin}}\, h_k \\ 0 \text{ otherwise} \end{cases} \tag{4.4.2}$$

The winning neuron updates its weight vector according to the learning rule:

$$\boldsymbol{w}_{k^*}^{new} = \boldsymbol{w}_{k^*}^{old} + \alpha(\boldsymbol{d}_{n.} - \boldsymbol{w}_{k^*}^{old}) \text{ where } \alpha \in [0,1] \text{ is the learning rate} \tag{4.4.3}$$

The characteristics of the data set $\boldsymbol{D}$ is then learnt by matching every input quadrixel to its closest weight vector in the sense of euclidean distance and this process iterates over and over until a termination decided by the operator.

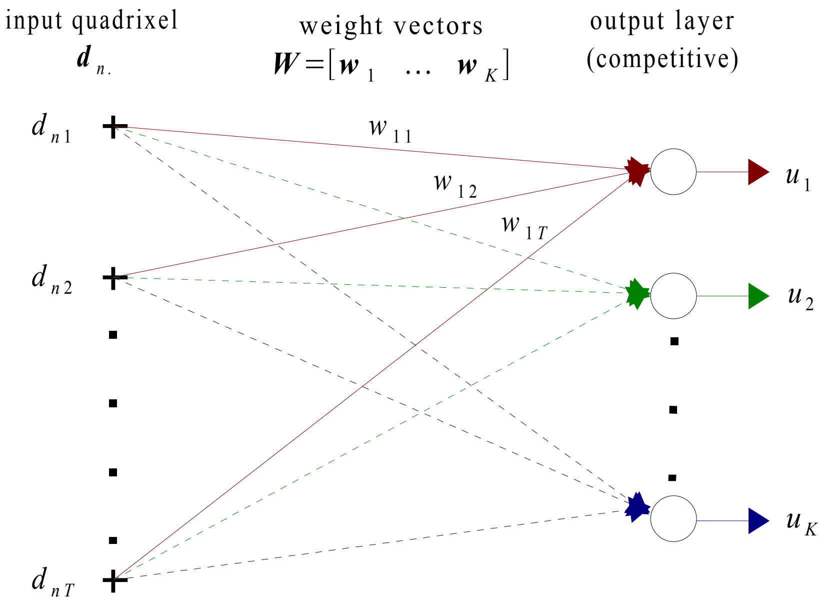


Figure 20: Kohonen Neural Network architecture

Kohonen learning rule

The network is trained from the input quadrixels chosen randomly one at a time.

Only the weights associated to the winning output neuron are updated. This is also known as a winner-take-all rule.

The Kohonen learning rule (4.4.3) can also be rewritten as

$$\boldsymbol{w}_{k^*}^{new} = (1-\alpha)\boldsymbol{w}_{k^*}^{old} + \alpha \boldsymbol{d}_n \tag{4.4.4}$$

(4.4.4) states that the learning rule moves the nearest weight vector towards the input.

Interpretation

At each presentation, the output neurons compete among themselves and the one whose weight vector is closest to the input quadrixel is identified (4.4.1) (4.4.2) and wins the privilege to learn: its weight vector is updated so that it is pushed toward the input quadrixel (4.4.3). Intuitively, it means that the best matching weight vector is adjusted to become even more similar to the input quadrixel. Consequently, the same neuron is more likely to win the competition each time a similar quadrixel is presented, but is less likely to win if a very different quadrixel is presented. As a result, after training, each weight vector will be adjusted toward a cluster of similar quadrixels. That is, all quadrixels that form a given cluster will always excite the same winning neuron that will output 1, while outputting 0 for all other different clusters. Thus, the Kohonen neural network is trained to recognize and label similar input quadrixels into clusters.

This type of learning strategy is partly inspired from observed brain activities in the cerebral cortex.

4.4.4 Kohonen Neural Network algorithm

$\boldsymbol{D} = \{\boldsymbol{d}_{1.}, \boldsymbol{d}_{2.}, \ldots, \boldsymbol{d}_{N.}\}$ is the given unlabelled data set of quadrixels.

Training process

1. Fix the number of clusters K, the number of epochs e_{max} and the learning rate $\alpha \in [0,1]$.

```
2. Initialize the weights $\boldsymbol{W}=[\boldsymbol{w}_1 \ \ \boldsymbol{w}_2 \ \ \dots \ \ \boldsymbol{w}_K]$.
3. For $e=1,\dots,e_{max}$
    ○ For $n=1,\dots,N$
        □ Select randomly a quadrixel $\boldsymbol{d}_{n.}$ from input data set $\boldsymbol{D}$.
        □ Competition. Find the winning neuron indexed by $k^*(\boldsymbol{d}_{n.})$:
          $k^*(\boldsymbol{d}_{n.})=\underset{k}{\operatorname{argmin}}\left\|\boldsymbol{d}_{n.}-\boldsymbol{w}_k\right\|$, $k=1\dots K$.
        □ Adaptation. Update the winner: $\boldsymbol{w}_{k^*}^{new}=\boldsymbol{w}_{k^*}^{old}+\alpha(\boldsymbol{d}_{n.}-\boldsymbol{w}_{k^*}^{old})$ (4.4.3).
        □ next $n$
    ○ next $e$
```

The number of epochs e_{max} is the number of times the entire data set $\boldsymbol{D}$ of quadrixels is presented to the neural network. A large number of epochs may improve the accuracy of the final weight vectors but will take more time.

The learning rate α can be viewed as how strong the nearest weight vectors are pushed toward their input quadrixel. A small learning rate may require more epochs during training, but a large learning rate may never converge. The value can range between 0 and 1. From (4.4.4), if $\alpha=0$, then the weight vectors are never updated. But if $\alpha=1$, then the weight vector $\boldsymbol{w}_{k^*}^{new}$ is made equal to the input quadrixel $\boldsymbol{d}_{n.}$.

Testing process

Segmentation is achieved by assigning voxels characterised by quadrixels $\boldsymbol{d}_{n.}$ to specific physiological regions c_k^* based on minimal distance criterion. That is, $\boldsymbol{d}_{n.}$ is assigned to cluster c_k for which the distance $h_k=\left\|\boldsymbol{w}_k-\boldsymbol{d}_{n.}\right\|$, $k=1\dots K$, is minimal. This can be done automatically by a testing process:

1. Each quadrixel $\boldsymbol{d}_{n.}$ is presented once more to the network which computes again the euclidean distances h_k (4.4.1) between $\boldsymbol{d}_{n.}$ and each of the final K weight vectors $\boldsymbol{W}$ found at the end of the training process.

2. A winning neuron $k^*(\boldsymbol{d}_{n.})$ is then identified and outputs $u_{k^*}=1$ (4.4.2) to automatically categorize the input quadrixel $\boldsymbol{d}_{n.}$: $\boldsymbol{d}_{n.}$ can be considered as belonging to the physiological region c_{k^*} whose representative TIC is given by the winning weight vector $\boldsymbol{w}_{k^*}=[w_{k^*1} \quad w_{k^*2} \quad \ldots \quad w_{k^*T}]'$.

4.4.5 Initialization problem and possible corrective actions

Initialization problem

If an output neuron has its weight vector initialized relatively far way from any input quadrixel, it might never win a competition and thus, might never be updated even if the training is continued.

Possible corrective actions

- Introduce cooperation between neurons:
  - During training, the weights of all losing neurons can also be updated, but at a smaller rate:

$$\begin{array}{l} \boldsymbol{w}_{k^*}^{new}=\boldsymbol{w}_{k^*}^{old}+\alpha_1(\boldsymbol{d}_n-\boldsymbol{w}_{k^*}^{old})\text{for winning neuron } k^* \\ \boldsymbol{w}_{k}^{new}=\boldsymbol{w}_{k}^{old}+\alpha_2(\boldsymbol{d}_n-\boldsymbol{w}_{k}^{old})\text{for all losing neurons } k\neq k^* \\ \text{with } \alpha_1 \gg \alpha_2 \end{array} \tag{4.4.5}$$

    (4.4.5) has the effect of slowly moving losing neuron weight vectors toward denser region of patterns.
  - The weights of the losing neurons which are in the neighbourhood of the winning neuron only can also be updated. This gives rise to the so-called Kohonen Self-Organizing Maps [32].
- Introduce a bias learning rule:

  A negative bias b_k is added to the euclidean distance (4.4.1) and updated at each iteration to help a neuron win more often. The value of the bias b_k depends on the frequency the corresponding neuron k wins. The bias is smaller for frequently winning neurons and larger for rarely winning neuron. Thus, a neuron that rarely wins has a large bias that helps push its weight vector to a cluster of input quadrixels. Then, the neuron can start to win more frequently and consequently, its bias can decrease to zero.

In this thesis, the Kohonen Neural Network is provided by the neural network toolbox 4.0 of MATLAB(R) 7.0, The MathWorks, Inc. which introduces this bias learning rule into its algorithm to overcome the weight initialization problem.

4.4.6 Termination problem

The main drawback of Kohonen Neural Network comes from the termination strategy. Indeed, the termination is artificially forced by the operator who fixes the number of epochs.

The learning rate $\alpha(e)$ could start with a large value and then be gradually reduced over the epochs to zero: this would force the algorithm to stop automatically when the weight vectors do not change very much any more. But yet, this learning strategy still does not guarantee a convergence of the weight vectors to any sensible model.

This is not the case for other clustering algorithms like K-means clustering which is based on an optimizing model and tries to minimize an objective function (4.2.1) reflecting the total squared error incurred in representing the data set $\boldsymbol{D}$ by the cluster centres $\boldsymbol{W}=[\boldsymbol{w}_1 \quad \boldsymbol{w}_2 \quad \dots \quad \boldsymbol{w}_K]$.

However, the great advantage of Kohonen Neural Network over K-means clustering is its ability to find clusters of heterogeneous shape and eventually, if the number of clusters is chosen properly, any group of similar TIC can be clustered.

4.5 Fuzzy K-Means Clustering (FKM)

4.5.1 Introduction

Hard clustering like Kohonen Neural Network assumes that the boundaries between clusters are well defined. So, each quadrixel, represented by a data point in the data space, is assigned to only one single cluster that corresponds to only one single physiological region. But due to noise and artefacts, like partial volume effect and breath motion on DCE-MR images, a more nuanced model to describe the uncertainty for a given quadrixel to belong to one single physiological region is desirable for clustering.

In this context, the so-called Fuzzy K-means clustering originally introduced by Bezdek [3] may be a more suitable clustering technique because it is assumed that each data point can belong to multiple clusters, but with varying degrees specified by membership functions.

Moreover, contrary to Kohonen neural network, Fuzzy K-means clustering is based on an optimizing model that minimizes a cost function by iteratively updating the cluster centres and the membership functions for each data point.

Wismuller *et al.* [67] performs a Fuzzy K-means clustering to discover clusters of similar temporal behaviours on dynamic cerebral contrast-enhanced perfusion MRI time-series. Chuang *et al.* [30] combines a Kohonen Neural Network and a Fuzzy K-means clustering to classify TIC of the brain according to temporal similarity in functional MR images. They especially show that the cascade clustering can overcome the difficulty to find activations occupying small areas of the cortex.

4.5.2 Cost Function minimization

Fuzzy K-means clustering seeks a minimum of this following cost function with respect to the membership functions $\boldsymbol{U}$ and the prototypes $\boldsymbol{W}$:

$$E_m(\boldsymbol{U},\boldsymbol{W})=\sum_{n=1}^{N}\sum_{k=1}^{K}u_{kn}^{m}\|\boldsymbol{d}_{n.}-\boldsymbol{w}_k\|^2 \qquad (4.5.1)$$

- $\boldsymbol{U}=\begin{bmatrix} u_{11} & \cdots & u_{1N} \\ \vdots & \ddots & \vdots \\ u_{K1} & \cdots & u_{KN} \end{bmatrix}$ is a fuzzy partition of the data set.

- u_{kn} is the membership function of $\boldsymbol{d}_{n.}$ to cluster c_k. A larger membership function u_{kn} indicates a higher confidence in the assignment of quadrixel $\boldsymbol{d}_{n.}$ to cluster c_k. It must satisfy the following two constraints:

$$0 \leqslant u_{kn} \leqslant 1 \ , \ \forall k = 1 \ldots K \text{ and } \forall n = 1 \ldots N \tag{4.5.2}$$

$$\sum_{k=1}^{K} u_{kn} = 1 \ , \ \forall n = 1 \ldots N \tag{4.5.3}$$

- $\boldsymbol{W} = [\boldsymbol{w}_1 \quad \boldsymbol{w}_2 \quad \ldots \quad \boldsymbol{w}_K]$ are the K prototypes of dimension T.
- $m > 1$ is a weighting exponent on each fuzzy membership that adjusts the degree of fuzziness between the different clusters. If $m = 0$, then the objective function becomes simply the sum-of-squared errors similar to (4.2.1) and each quadrixel is assigned to only one cluster.

The objective function (4.5.1) to be minimized can be intuitively interpreted as the sum of the squared distances in the euclidean space between all the quadrixels and the cluster centres weighted by the corresponding membership functions.

4.5.3 Fuzzy K-Means algorithm

Given the unlabelled data set $\boldsymbol{D} = \{\boldsymbol{d}_{1.}, \boldsymbol{d}_{2.}, \ldots, \boldsymbol{d}_{N.}\}$, the fuzzy clustering is carried out through an iterative process to optimize $E_m(\boldsymbol{U}, \boldsymbol{W})$ (4.5.1) and obtain a fuzzy partition $\boldsymbol{U}$ of $\boldsymbol{D}$:

1. Fix the number of clusters K, the weighting coefficient $m > 1$ and the membership functions $\boldsymbol{U}$ such that (4.5.2) and (4.5.3) hold.
2. Update the prototypes $\boldsymbol{W} = [\boldsymbol{w}_1 \quad \boldsymbol{w}_2 \quad \ldots \quad \boldsymbol{w}_K]$ for all clusters c_k, $k = 1 \ldots K$:

$$\boldsymbol{w}_k = \frac{\sum_{n=1}^{N} u_{kn}^{m} \boldsymbol{d}_n}{\sum_{n=1}^{N} u_{kn}^{m}} \tag{4.5.4}$$

3. Update the membership functions u_{kn} for all quadrixels $\boldsymbol{d}_{n.}$, $n = 1 \ldots N$ and for all clusters c_k, $k = 1 \ldots K$:

$$u_{kn} = \frac{1}{\sum_{l=1}^{K} \left(\frac{\|\boldsymbol{d}_{n.} - \boldsymbol{w}_k\|}{\|\boldsymbol{d}_{n.} - \boldsymbol{w}_l\|} \right)^{\frac{2}{m-1}}} \tag{4.5.5}$$

4. if the change between two consecutive prototypes $\boldsymbol{W}$ or membership functions $\boldsymbol{U}$ is less than a given threshold, then stop the procedure. Else return to 2 .

Intuitively, a given membership function (4.5.5) measures the similarity between the corresponding quadrixel and cluster centre: the closer they are in the sense of Euclidean distance, the higher the membership function.

Clearly, the Hard K-means clustering is only a special case where the membership functions are given by

$$u_{k^* n} = \begin{cases} 1 & \text{if } k^* = \underset{k}{\text{argmin}} \|\boldsymbol{d}_{n.} - \boldsymbol{w}_k\| \\ 0 & \text{otherwise} \end{cases} \tag{4.5.6}$$

Fuzzy K-means clustering is better than the Hard K-means clustering at avoiding local minima but can still converge to local minima of the cost function given by (4.5.1).

The main advantage of Fuzzy K-means clustering is its ability to group quadrixels into multiple clusters with varying degrees of membership and thus, takes into account that boundaries have no sharp transitions because of noise, artefacts and overlaps that occur in MR studies.

This is particularly useful for voxels which are hard to label automatically based on their TIC. In this case, assigning the quadrixel to the cluster with the highest membership value may not always be the best approach. Instead, displaying the membership functions as K volumes that reflect the degree of confidence for each quadrixel to belong to any of the K clusters produces a fuzzy map that retains more information about uncertain quadrixels. Then, an operator can select a reasonable tissue membership value based on visual inspection to threshold the fuzzy map and thus obtain a hard clustering of the data set.

4.6 Expectation-Maximization Clustering (EM)

4.6.1 Introduction

Expectation-Maximization clustering (EM) [1] is a statistical approach to group similar patterns of data.

Even though quadrixels of a given physiological and anatomical region are characterised by a common pattern of TIC, they do not share the same exact values over time. That is, when mapped into a high-dimensional space, they do not plot into a same single point but rather into multiple distinct points that can be grouped into a cluster c_k. The shape and the size of cluster c_k reflect the distribution of the quadrixel values around a certain mean $\boldsymbol{w}_k$ which is located at the centre of the cluster.

While Hard K-means clustering, Kohonen neural network and Fuzzy K-means clustering all try to characterise a cluster c_k with a mean value $\boldsymbol{w}_k$ only, EM clustering defines it with a probability density function that must be learnt in an iterative process from the data set.

Thus, to express the variability of the quadrixel values in probabilistic terms, $\boldsymbol{d}$ is considered to be a continuous random variable whose distribution $p(\boldsymbol{d}|c_k)$ depends on the type of physiological and anatomical region c_k it belongs to.

When all the relevant probability values are learnt from EM, hard clustering can be performed subsequently by a maximum a posteriori classifier.

4.6.2 The Gaussian mixture model

In a Gaussian mixture model, it is assumed that the quadrixels belonging to each of the K physiological regions are distributed according to respectively K Gaussian distributions $p(\boldsymbol{d}|c_k)$ of parameters $\{\boldsymbol{w}_k, \boldsymbol{\Sigma}_k\}$, $k=1\ldots K$, where $\boldsymbol{w}_k$ is the mean and $\boldsymbol{\Sigma}_k$ the covariance matrix:

$$p(\boldsymbol{d}|c_k)=\frac{1}{(2\pi)^{\frac{T}{2}}\sqrt{|\boldsymbol{\Sigma}_k|}}\exp\left(-\frac{1}{2}(\boldsymbol{d}-\boldsymbol{w}_k)'\boldsymbol{\Sigma}_k^{-1}(\boldsymbol{d}-\boldsymbol{w}_k)\right),\ k=1\ldots K \tag{4.6.1}$$

Thus, a Gaussian mixture model refines the model implied by previous algorithms by parametrizing clusters with true probability density functions $p(\boldsymbol{d}|c_k)$ described not only by means $\boldsymbol{w}_k$ but also associated covariance matrices $\boldsymbol{\Sigma}_k$, $k=1\ldots K$.

Consequently, the probability density function of the whole data set is given by a mixture density:

$$p(\boldsymbol{d})=\sum_{k=1}^{K} P(c_k)\,p(\boldsymbol{d}|c_k) \tag{4.6.2}$$

with

$$0\leqslant P(c_k)\leqslant 1 \text{ and } \sum_{k=1}^{K} P(c_k)=1 \tag{4.6.3}$$

The previous equations can be interpreted as follows:

- (4.6.1) is a Gaussian distribution whose parameters are $\{\boldsymbol{w}_k , \boldsymbol{\Sigma}_k\}$ and describes the variability of quadrixels values in the physiological region c_k . It can be regarded as a “sub-model”.
- (4.6.2) states that the density of the data set $\boldsymbol{d}$ can be fitted by a “super-model” which is a mixture of K individual Gaussian distributions $p(\boldsymbol{d}|c_k)$ (4.6.1) weighted by respective coefficients $P(c_k)$ (4.6.3).
- (4.6.3) is a mixture weight that stands for the relative importance of each sub-model in the super-model.

Intuitively, combining several sub-models to produce a super-model is often useful to best fit the density of the data set, especially in case of MRU studies: a given quadrixel may be a mixture of multiple TIC and thus, may not be sufficiently explained by the Gaussian distribution of one single distinct cluster when mapped into the data space.

Moreover, it is commonly accepted that any continuous data density can be approximated relatively well by Gaussian mixture models with suitable parameters $\{\boldsymbol{w}_k , \boldsymbol{\Sigma}_k\}$, mixture weights $P(c_k)$ and number of sub-models K .

So, given unlabelled data set $\boldsymbol{D}=\{\boldsymbol{d}_1 , \boldsymbol{d}_2 , \ldots , \boldsymbol{d}_N\}$ and the number of clusters K , the aim of Expectation-Maximization algorithm is to estimate the parameters $\{\hat{\boldsymbol{w}}_k , \hat{\boldsymbol{\Sigma}}_k\}$,

$k=1\ldots K$, of the K Gaussian distributed clusters and their respective weights $\hat{P}(c_k)$ that best describe the density of the data set $\boldsymbol{D}$ in the T dimensional space.

Put it in another way, it can be considered that all the data points $\boldsymbol{d}_{n.}$, $n=1\ldots N$, were generated from the probability density function given by (4.6.2) as follows:

- A cluster c_k is randomly chosen with probability $P(c_k)$.
- A data point $\boldsymbol{d}_{n.}$ is generated from the corresponding Gaussian distribution $p(\boldsymbol{d}|c_k)$ (4.6.1).
- The process is repeated over for the N data points.

The aim of EM can then be interpreted as finding the parameters $\hat{\boldsymbol{w}}_k$, $\hat{\boldsymbol{\Sigma}}_k$ and $\hat{P}(c_k)$, $k=1\ldots K$, that are most likely to generate the observed data points $\boldsymbol{D}=\{\boldsymbol{d}_{1.},\boldsymbol{d}_{2.},\ldots,\boldsymbol{d}_{N.}\}$.

4.6.3 Maximum-Likelihood estimation

Likelihood

It is assumed that the unlabelled quadrixels $\boldsymbol{d}_{n.}$, $n=1\ldots N$, are drawn independently from the Gaussian mixture density specified by (4.6.2):

$$p(\boldsymbol{d}_{n.})=\sum_{k=1}^{K} p(\boldsymbol{d}_{n.}|c_k)P(c_k) \tag{4.6.4}$$

So, the resulting density of the entire data set $\boldsymbol{D}=\{\boldsymbol{d}_{1.},\boldsymbol{d}_{2.},\ldots,\boldsymbol{d}_{N.}\}$, called likelihood, can be written as:

$$p(\boldsymbol{D})=\prod_{n=1}^{N} p(\boldsymbol{d}_{n.}) \tag{4.6.5}$$

The strategy employed by EM is to find the Gaussian mixture that best fits the data set $\boldsymbol{D}$ by estimating the parameters $\hat{\boldsymbol{\theta}}_k=\{\hat{\boldsymbol{w}}_k,\hat{\boldsymbol{\Sigma}}_k,\hat{P}(c_k)\}$, $k=1\ldots K$, that maximize the likelihood (4.6.5). Thus, $\hat{\boldsymbol{\theta}}_k=\{\hat{\boldsymbol{w}}_k,\hat{\boldsymbol{\Sigma}}_k,\hat{P}(c_k)\}$, $k=1\ldots K$, are called maximum-likelihood estimates.

Usually, the log-likelihood L of the entire data set is rather considered:

$$L = \log p(\boldsymbol{D}) = \log \prod_{n=1}^{N} p(\boldsymbol{d}_{n.}) = \sum_{n=1}^{N} \log p(\boldsymbol{d}_{n.}) = \sum_{n=1}^{N} \log \left\{ \sum_{k=1}^{K} P(c_k)\, p(\boldsymbol{d}_{n.}|c_k) \right\} \tag{4.6.6}$$

Posterior probability or responsibility

L must be maximized with respect to $\boldsymbol{\theta}_k = \{\boldsymbol{w}_k, \boldsymbol{\Sigma}_k, P(c_k)\}$ and subject to the constraints given by (4.6.3).

For that, the posterior probability $P(c_k|\boldsymbol{d}_{n.})$ is introduced according to the Bayes formula:

$$P(c_k|\boldsymbol{d}_{n.}) = \frac{p(\boldsymbol{d}_{n.}|c_k)\, P(c_k)}{p(\boldsymbol{d}_{n.})} \tag{4.6.7}$$

In the light of (4.6.7), $P(c_k)$, $p(\boldsymbol{d}_{n.}|c_k)$, $P(c_k|\boldsymbol{d}_{n.})$ and $p(\boldsymbol{d}_{n.})$ can be reinterpreted as follow:

- $P(c_k)$ is the prior probability that quadrixel $\boldsymbol{d}_{n.}$ belongs to region c_k. This probability reflects the prior knowledge of how likely $\boldsymbol{d}_{n.}$ belongs to c_k before it is observed
- $p(\boldsymbol{d}_{n.}|c_k)$ is a class-conditional probability density function that shows the probability density of observing a particular quadrixel value $\boldsymbol{d}_{n.}$ given that it belongs to region c_k
- $P(c_k|\boldsymbol{d}_{n.})$ is the posterior probability that $\boldsymbol{d}_{n.}$ is assigned to region c_k given that the quadrixel value is observed to be $\boldsymbol{d}_{n.}$. It shall be called responsibility of the cluster c_k for data point $\boldsymbol{d}_{n.}$.
- $p(\boldsymbol{d}_{n.})$ is a scale factor that guarantees $\sum_{k=1}^{K} P(c_k|\boldsymbol{d}_{n.}) = 1$

The Bayes formula (4.6.7) merely states that the observation of the value of quadrixel $\boldsymbol{d}_{n.}$ converts the prior probability $P(c_k)$ to the posterior probability $P(c_k|\boldsymbol{d}_{n.})$.

The responsibility $P(c_k|\boldsymbol{d}_{n.})$ plays a major role since it gives the probability membership of quadrixel $\boldsymbol{d}_{n.}$ to a physiological region c_k. If a hard segmentation has to be made, the maximum a posteriori (MAP) classifier can be used: quadrixel $\boldsymbol{d}_{n.}$ is assigned to the cluster c_{k^*} with the highest responsibility.

Maximum-Likelihood estimates

Moreover, responsibility $P(c_k|\boldsymbol{d}_{n.})$ is also involved in computing the maximum-likelihood estimates which maximize L and which are given, without demonstrations, as follows:

$$\hat{P}(c_k)=\frac{1}{N}\sum_{n=1}^{N} P(c_k|\boldsymbol{d}_{n.}) \tag{4.6.8}$$

$$\hat{\boldsymbol{w}}_k=\frac{\sum_{n=1}^{N} P(c_k|\boldsymbol{d}_{n.})\,\boldsymbol{d}_{n.}}{\sum_{n=1}^{N} P(c_k|\boldsymbol{d}_{n.})} \tag{4.6.9}$$

$$\hat{\boldsymbol{\Sigma}}_k=\frac{\sum_{n=1}^{N} P(c_k|\boldsymbol{d}_n)(\boldsymbol{d}_n-\hat{\boldsymbol{w}}_k)(\boldsymbol{d}_n-\hat{\boldsymbol{w}}_k)'}{\sum_{n=1}^{N} P(c_k|\boldsymbol{d}_n)} \tag{4.6.10}$$

The maximum-likelihood estimates have a very intuitive interpretation:

- (4.6.8) states that the relative importance of the sub-model cluster c_k in the mixture super-model is given by the average posterior probability that each of the data points $\boldsymbol{d}_{n.}$, $n=1\ldots N$, belongs to cluster c_k. That is, the more likely the data points to belong to cluster c_k, the larger cluster weight $P(c_k)$ to describe the super-model.
- (4.6.9) states that the mean $\boldsymbol{w}_k$ of cluster c_k is given by the average data points $\boldsymbol{d}_{n.}$, $n=1\ldots N$, weighted by the posterior probability that each of the data points belongs to cluster c_k.
- (4.6.10) states that the covariance matrix $\boldsymbol{\Sigma}_k$ of cluster c_k is given by the variance of the data points $\boldsymbol{d}_{n.}$, $n=1\ldots N$, with respect to the mean $\boldsymbol{w}_k$, weighted by the posterior probability that each of the data points belongs to cluster c_k.

4.6.4 EM algorithm

Unfortunately, (4.6.8), (4.6.9) and (4.6.10) do not give explicitly $\hat{\boldsymbol{\theta}}_k=\left\{\hat{\boldsymbol{w}}_k, \hat{\boldsymbol{\Sigma}}_k, \hat{P}(c_k)\right\}$, $k=1\ldots K$, but rather a coupled set of non-linear simultaneous equations. So, given the unlabelled data set $\boldsymbol{D}=\{\boldsymbol{d}_{1.},\boldsymbol{d}_{2.},\ldots,\boldsymbol{d}_{N.}\}$, the most obvious approach to solve for $\hat{\boldsymbol{\theta}}_k$ is an iterative procedure that involves applying alternatively a E-step and a M-step:

1. Fix the number of clusters K, make an initial guess for the parameters $\{\hat{\boldsymbol{w}}_k,\hat{\boldsymbol{\Sigma}}_k\}$ of the K Gaussian distributions and for the mixture weights $\hat{P}(c_k)$, $k=1\ldots K$.
2. **E-step**: compute the responsibility $\hat{P}(c_k|\boldsymbol{d}_{n.})$ of each cluster c_k, $k=1\ldots K$, for each data point $\boldsymbol{d}_{n.}$, $n=1\ldots N$ given by
   - ○ $\hat{P}(c_k|\boldsymbol{d}_{n.})=\dfrac{\hat{p}(\boldsymbol{d}_{n.}|c_k)\hat{P}(c_k)}{\hat{p}(\boldsymbol{d}_{n.})}$ (refer to 4.6.7)
   - ○ $\hat{p}(\boldsymbol{d}_{n.}|c_k)=\dfrac{1}{(2\pi)^{\frac{T}{2}}\sqrt{|\hat{\boldsymbol{\Sigma}}_k|}}\exp\left(-\dfrac{1}{2}(\boldsymbol{d}_{n.}-\hat{\boldsymbol{w}}_k)'\hat{\boldsymbol{\Sigma}}_k^{-1}(\boldsymbol{d}_{n.}-\hat{\boldsymbol{w}}_k)\right)$ (refer to 4.6.1)
   - ○ $\hat{p}(\boldsymbol{d}_{n.})=\sum_{k=1}^{K}\hat{p}(\boldsymbol{d}_{n.}|c_k)\hat{P}(c_k)$ (refer to 4.6.4)
3. **M-step**: Re-estimate the parameters $\{\hat{\boldsymbol{w}}_k,\hat{\boldsymbol{\Sigma}}_k\}$ of the K Gaussian distributions and the mixture weights $\hat{P}(c_k)$, $k=1\ldots K$, based on an estimate of the responsibilities $\hat{P}(c_k|\boldsymbol{d}_{n.})$ and data points $\boldsymbol{d}_{n.}$:
   - ○ $\hat{P}(c_k)=\dfrac{1}{N}\sum_{n=1}^{N}\hat{P}(c_k|\boldsymbol{d}_{n.})$ (refer to 4.6.8)
   - ○ $\hat{\boldsymbol{w}}_k=\dfrac{\sum_{n=1}^{N}\hat{P}(c_k|\boldsymbol{d}_{n.})\,\boldsymbol{d}_{n.}}{\sum_{n=1}^{N}\hat{P}(c_k|\boldsymbol{d}_{n.})}$ (refer to 4.6.9)
   - ○ $\hat{\boldsymbol{\Sigma}}_k=\dfrac{\sum_{n=1}^{N}\hat{P}(c_k|\boldsymbol{d}_n)(\boldsymbol{d}_n-\hat{\boldsymbol{w}}_k)(\boldsymbol{d}_n-\hat{\boldsymbol{w}}_k)'}{\sum_{n=1}^{N}\hat{P}(c_k|\boldsymbol{d}_n)}$ (refer to 4.6.10)
4. If the change between two consecutive log-likelihoods L (4.6.6) is less than a given threshold, then stop the procedure. Else return to 2.

An interesting property of EM is that the log-likelihood can only increase over the iterations. Yet, there is no guarantee the algorithm necessarily converges to a global maximum of the log-likelihood. So, the results are highly dependant on the initialization of the parameters $\{\hat{\boldsymbol{w}}_k, \hat{\boldsymbol{\Sigma}}_k\}$ and the mixture weights $\hat{P}(c_k)$, $k=1\ldots K$.

A standard approach is to preliminary cluster the data points $\boldsymbol{D}$ by a hard K-means clustering to get good starting values:

- the means $\hat{\boldsymbol{w}}_k$ are initialized at the centre of the clusters c_k
- the covariance matrices $\hat{\boldsymbol{\Sigma}}_k$ are set equal to $\frac{1}{N_k-1}\sum_{\boldsymbol{d}_n\in c_k}(\boldsymbol{d}_n-\boldsymbol{w}_k)(\boldsymbol{d}_n-\boldsymbol{w}_k)'$ where N_k is the number of data points $\boldsymbol{d}_n$ that belong to cluster c_k
- the mixture weights $\hat{P}(c_k)$ are chosen equal to $\frac{N_k}{N}$

4.6.5 Dimensionality reduction

EM clustering is computationally intensive, especially because it involves at each iteration the computation of the inverse and the determinant of K covariance matrices $\boldsymbol{\Sigma}_k$ (4.6.1) in a T-dimensional space where $24<T<44$ in the case of our MRU studies.

The EM clustering can be faster and use less memory if an appropriate subspace of dimension $T_2\leq T$ is determined by a principal component analysis (PCA) or Karhunen-Loève transform [35]:

$$\boldsymbol{D}_2=\boldsymbol{D}\,\boldsymbol{A} \tag{4.6.11}$$

- $\boldsymbol{D}$ is the $(N x T)$ data set matrix
- $\boldsymbol{D}_2$ is the $(N\times T_2)$ derived data set matrix
- $\boldsymbol{A}$ is the $(T\times T_2)$ linear transformation matrix whose columns are the first T_2 eigenvectors corresponding to the first T_2 largest eigenvalues of the covariance matrix $\frac{1}{N-1}\boldsymbol{D}'\boldsymbol{D}$ of the data set $\boldsymbol{D}$.

$\boldsymbol{A}$ is known as the similarity transform. It retains from the data set $\boldsymbol{D}$, only the $T_2<T$ transformed features which are uncorrelated with each other and have the largest variances.

It can be shown that $\boldsymbol{D}_2$ optimally approximates in the least squares sense the data set $\boldsymbol{D}$ in a lower dimensional sub-space $T_2<T$ and thus, EM would rather cluster $\boldsymbol{D}_2$

instead of the original data set $\boldsymbol{D}$ to accelerate the algorithm without any significant loss of variance.

The EM algorithm is implemented to include a PCA whose T_2 value is determined by retaining 95% of the variance.

However, it shall be noted that the dimensionality reduction by PCA is suitable for faithful representation of the data set, not for discrimination of patterns. Indeed, emphasis is placed on features with greatest variability rather than the ones that can separate the clusters.

4.6.6 Summary of EM properties

EM clustering has the advantage over all the previous clustering methods that it can detect clusters of very different size and shape. Basically, the number of data points N_k which are likely to belong to cluster c_k is controlled by the mixture weight $P(c_k)$ and the shape of the cluster is controlled by the covariance matrix $\boldsymbol{\Sigma}_k$.

Moreover, it also has the same interesting properties as Fuzzy K-means: quadrixels are grouped into multiple clusters with varying degrees of membership given by posterior probabilities. Thus, probability maps that reflect how likely each quadrixel belongs to any of the K clusters can be displayed to an operator who selects a reasonable tissue membership threshold to get a final hard clustering of the data set.

However, EM clustering model is entirely based on the assumption that quadrixel distribution in MRU studies can be expressed as a mixture of Gaussian distributions, which is not necessarily true, especially for noisy data sets.

4.7 Factor Analysis of Medical Image Series (FAMIS)

4.7.1 Introduction

Factor Analysis of Medical Image Series (FAMIS) extracts, from dynamic image series, physiological regions related to anatomical or functional structures with different temporal behaviours, even if they partially overlap and yield mixed kinetics.

The method was primarily developed for nuclear medicine data [2] [5] [6] [8] [10] [15] [26] [28] [46] as a good alternative to the widely used manual region of interest approach. But it has also been applied more recently on MRI [13] [21] [41] [42] [44] [49] and other imaging modalities (PET [17] [22] [58], SPECT [33] [36] [45], ultrasound [55], spectral sequences [29] [56]) and extended to process dynamic volumetric data sets [13] [17] [58]. The applications are very various (brain studies [2] [17] [41] [42], cardiac studies [6] [22] [28] [36] [45] [44] [55] [58], renal studies [8] [26] [46], hepatic studies [15], angiography [49], oncology [13] [21], noise removal [28]).

FAMIS is based on the assumption that the relevant information in the dynamic image series can be summarized into a limited number K of fundamental spatial distributions, called factor volumes, wherein the fundamental TIC, called associated factor curve, is homogeneous (Figure 21). But volumes may also partially overlap due to partial volume effects for example. Therefore, the TIC measured at each voxel of the studied image sequence, called a quadrixel, is considered as a composite TIC that can be decomposed into a linear combination of the K underlying factor curves:

$$d_{nt} = \sum_{k=1}^{K} a_{nk} f_{kt} + \varepsilon_{nt} \tag{4.7.1}$$

- K is the number of underlying structures or physiological regions
- d_{nt} is one TIC sample of voxel n at time point t
- a_{nk} is a time-independent factor coefficient that represents a weight: it expresses the contribution of factor curve f_{kt} in the composite TIC d_{nt} and differs from one quadrixel to another; the set of factor coefficients $\boldsymbol{a}_{.k} = [a_{1\mathrm{k}} \dots a_{nk} \dots a_{Nk}]'$, called factor volume, corresponds to the fundamental spatial distribution of a physiological region

- $f_{k\,t}$ is the k^{th} factor curve at time point t; a factor curve $\boldsymbol{f}_{k\,.}$ is the fundamental TIC of a physiological region
- $\varepsilon_{n\,t}$ is the error due to both the model and the noise.

For demonstration purposes, it is convenient to consider that in our case of MRU studies, a quadrixel $\boldsymbol{d}_{n.}$ that enhances may be expressed for example as:

$$\boldsymbol{d}_{n\,.} = a_{n\,1} \cdot \boldsymbol{f}_{1\,.} + a_{n\,2} \cdot \boldsymbol{f}_{2\,.} + a_{n\,3} \cdot \boldsymbol{f}_{3\,.} + \boldsymbol{\varepsilon}_{n\,.} \text{ with } K\text{=3} \tag{4.7.2}$$

where $\boldsymbol{f}_{1\,.}$, $\boldsymbol{f}_{2\,.}$, and $\boldsymbol{f}_{3\,.}$ are the factor curves and represent the fundamental TIC of a pure voxel that would only belong respectively to the kidney, the liver and the spleen.

Factor curves and their associated factor volumes can be interpreted in a symmetric or dual way: one can also consider that a volume image at time point t in the sequence may also be expressed as:

$$\boldsymbol{d}_{.\,t} = f_{1\,\mathrm{t}} \cdot \boldsymbol{a}_{.\,1} + f_{2\,\mathrm{t}} \cdot \boldsymbol{a}_{.\,2} + f_{3\,\mathrm{t}} \cdot \boldsymbol{a}_{.\,3} + \boldsymbol{\varepsilon}_{.\,t} \text{ with } K\text{=3} \tag{4.7.3}$$

(4.7.2) states that one quadrixel $\boldsymbol{d}_{n.}$ can be decomposed over $K=3$ factor curves weighted by their associated factor coefficients. (4.7.3) states that one volume image $\boldsymbol{d}_{.t}$ at time point t can be decomposed over $K=3$ factor volumes weighted by their associated factor curves at time point t.

The weight $a_{n\,k}$ associated to factor curve $\boldsymbol{f}_{k\,.} = \left[f_{k\,1} \ldots f_{k\,t} \ldots f_{k\,T} \right]'$ is given by the voxel intensity in the associated factor volume $\boldsymbol{a}_{.\,k} = \left[a_{1\,\mathrm{k}} \ldots a_{n\,k} \ldots a_{N\,k} \right]'$, and hence, each factor volume represents the spatial distribution of the associated factor curve $\boldsymbol{f}_{k\,.}$.

It must be noticed right now that, in cases where TIC are considered, $\forall k=1 \ldots K$ and $\forall t=1 \ldots T$, f_{kt} cannot be by nature negative as they represent fundamental signal intensities over time, and symmetrically, $\forall n=1 \ldots N$ and $\forall k=1 \ldots K$, a_{nk} must also be non negative as they represent voxel intensities at a given time.

The aim of FAMIS is then to find the physiological factor volumes and factor curves that linearly decompose the studied image sequence, assuming that K is known. It generally includes two main stages:

- **Orthogonal analysis.** The orthogonal analysis is merely a principal component analysis (PCA) or a singular value decomposition (SVD). It determines a lower dimensional orthogonal subspace, also called 'study space', that is optimal in the least squares sense and separates most of the relevant part of the quadrixels from the noise.

- **Oblique analysis.** The basis vectors of the study subspace found by the orthogonal analysis cannot represent the physiological factor curves. Indeed, they are by definition orthogonal and thus necessarily contain negative values. Consequently, an oblique rotation under constraints is performed to obtain non-orthogonal basis vectors, namely the factor curves.

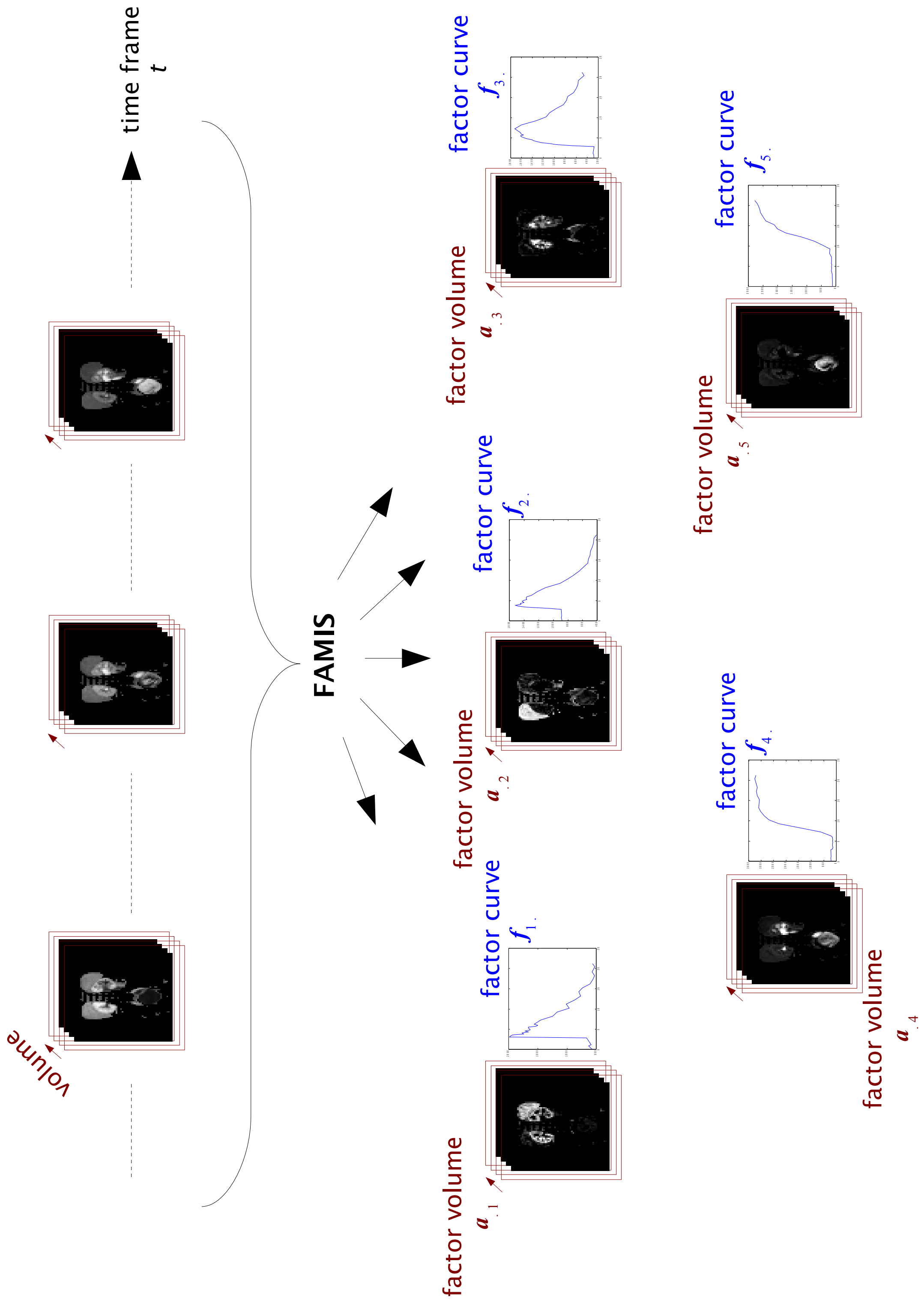


Figure 21: Decomposition into factor curves and factor volumes

4.7.2 FAMIS model and previous approaches

FAMIS model

To recall, after preprocessing, a dynamic volumetric data set $\boldsymbol{D}$ is considered as a set of N enhanced quadrixels $\boldsymbol{d}_{n\,.}$, $n=1\ldots N$, of T components d_{nt}, $t=1\ldots T$, where T is the number of time points.

All the information in the data set $\boldsymbol{D}$ is contained in these N quadrixels and from (4.7.1), it can also be expressed as

$$\boldsymbol{D} = \boldsymbol{A}\,\boldsymbol{F} + \boldsymbol{E}_1 \tag{4.7.4}$$

- $\boldsymbol{D}=\begin{bmatrix} d_{11} & \cdots & d_{1T} \\ \vdots & \ddots & \vdots \\ d_{N1} & \cdots & d_{NT} \end{bmatrix}_{(N\times T)} = \begin{bmatrix} \boldsymbol{d}_{1\,.}{}' \\ \vdots \\ \boldsymbol{d}_{N\,.}{}' \end{bmatrix} = \begin{bmatrix} \boldsymbol{d}_{.\,1} & \cdots & \boldsymbol{d}_{.\,T} \end{bmatrix}$

- $\boldsymbol{A}=\begin{bmatrix} a_{11} & \cdots & a_{1K} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NK} \end{bmatrix}_{(N\times K)} = \begin{bmatrix} \boldsymbol{a}_{.\,1} & \cdots & \boldsymbol{a}_{.\,K} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_{1\,.}{}' \\ \vdots \\ \boldsymbol{a}_{N\,.}{}' \end{bmatrix}$

- $\boldsymbol{F}=\begin{bmatrix} f_{11} & \cdots & f_{1T} \\ \vdots & \ddots & \vdots \\ f_{K1} & \cdots & f_{KT} \end{bmatrix}_{(K\times T)} = \begin{bmatrix} \boldsymbol{f}_{1\,.}{}' \\ \vdots \\ \boldsymbol{f}_{K\,.}{}' \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}_{.\,1} & \cdots & \boldsymbol{f}_{.\,T} \end{bmatrix}$

- $\boldsymbol{E}_1=\begin{bmatrix} \varepsilon_{11} & \cdots & \varepsilon_{1T} \\ \vdots & \ddots & \vdots \\ \varepsilon_{N1} & \cdots & \varepsilon_{NT} \end{bmatrix}_{(N\times T)}$

A is the matrix whose columns are the factor volumes and F is the matrix whose rows are the associated factor curves. The spatial distribution of factor curve $\boldsymbol{f}_{k\,.}=\left[f_{k1}\ldots f_{kt}\ldots f_{kT}\right]'$ is contained in the factor volume $\boldsymbol{a}_{.\,k}=\left[a_{1k}\ldots a_{nk}\ldots a_{Nk}\right]'$.

The goal of FAMIS is to extract the matrices $\boldsymbol{A}$ and $\boldsymbol{F}$ that have physiological meanings.

Orthogonal analysis. The matrix $\boldsymbol{D}$ is factorized by extracting the eigenvectors of the covariance matrix $\boldsymbol{C}=\frac{1}{N}\boldsymbol{D}'\boldsymbol{D}$ and by keeping only the first K eigenvalues that account for most of the variance:

$$\boldsymbol{D} = \boldsymbol{V}\ \boldsymbol{\Lambda}\ \boldsymbol{U} + \boldsymbol{E}_2 \tag{4.7.5}$$

- $\boldsymbol{U}$ is the $(K \times T)$ matrix of the first K eigenvectors of $\boldsymbol{C}$ and $\boldsymbol{U}\,\boldsymbol{U}' = \boldsymbol{I}_K$ where $\boldsymbol{I}_K$ is the $(K \times K)$ identity matrix.
- $\boldsymbol{\Lambda}$ is the $(K \times K)$ diagonal matrix of the square roots of the first K eigenvalues.
- $\boldsymbol{V} = \boldsymbol{D}\,\boldsymbol{U}'\boldsymbol{\Lambda}^{-1}$ is the $(N \times K)$ matrix of the coordinates of the quadrixels in the subspace spanned by the first K eigenvectors and $\frac{1}{N}\boldsymbol{V}\ \boldsymbol{V}' = \boldsymbol{I}_N$

Substituting for $\boldsymbol{D}$ from (4.7.4) and (4.7.5) and ignoring the error matrices:

$$\boldsymbol{A} = \boldsymbol{V}\ \boldsymbol{\Lambda}\ \boldsymbol{R}^{-1} \tag{4.7.6}$$

$$\boldsymbol{F} = \boldsymbol{R}\ \boldsymbol{U} \tag{4.7.7}$$

where $\boldsymbol{R}$ is a non singular $(K \times K)$ matrix with $\boldsymbol{R}^{-1}\boldsymbol{R} = \boldsymbol{I}_K$.

$\boldsymbol{A} = \boldsymbol{V}\ \boldsymbol{\Lambda}$ and $\boldsymbol{F} = \boldsymbol{U}$ with $\boldsymbol{R} = \boldsymbol{I}_K$ are trivial solutions that are of no interest because they are two orthogonal matrices and thus necessarily contain negative values.

Oblique analysis. This stage is necessary to perform a rotation $\boldsymbol{R} \neq \boldsymbol{I}_K$ on the principal components $\boldsymbol{U}$ so that the factors $\boldsymbol{A}$ and $\boldsymbol{F}$ can represent underlying physiology. The most common approach is an iterative procedure that takes into account the positivity constraints on $\boldsymbol{A}$ and $\boldsymbol{F}$.

Starting from a first estimate of $\boldsymbol{R}$, the factor volumes $\boldsymbol{A}$ can be computed from (4.7.6). Then, $\boldsymbol{A}$ is modified to satisfy the positivity constraints.

Multiplying both sides of (4.7.6) on the left by $(\boldsymbol{V}'\boldsymbol{V})^{-1}\boldsymbol{V}'$, a new rotation matrix $\boldsymbol{R}$ can be computed from the modified $\boldsymbol{A}$:

$$\boldsymbol{R}^{-1} = \boldsymbol{\Lambda}^{-1}(\boldsymbol{V}'\boldsymbol{V})^{-1}\boldsymbol{V}'\boldsymbol{A} \tag{4.7.8}$$

From (4.7.7), $\boldsymbol{F}$ is obtained from the new $\boldsymbol{R}$ and then is also modified to satisfy the positivity constraints.

Both sides of (4.7.7) are multiplied on the right by $\boldsymbol{U}'(\boldsymbol{U}\,\boldsymbol{U}')^{-1}$ and a new rotation matrix $\boldsymbol{R}$ again can be recomputed from the modified $\boldsymbol{F}$:

$$\boldsymbol{R} = \boldsymbol{F}\,\boldsymbol{U}'(\boldsymbol{U}\,\boldsymbol{U}')^{-1} \tag{4.7.9}$$

Then A is recomputed from (4.7.6) again and this is performed iteratively to satisfy the positivity constraints until convergence.

However, positivity constraints alone are usually not sufficient to produce an unique solution [7]. This is the main problem of FAMIS and it has given rise to many different approaches.

Previous approaches

Several groups [2] [5] [6] [13] [21] [22] [56] propose an iterative apex seeking procedure: the trixels are normalized, centred and projected onto a $K-1$ dimensional subspace spanned by the first $K-1$ principal components. Then, the apices of a polytope that contain all the trixel projection are isolated under positivity constraints to estimate the physiological factor curves. Buvat *et al.* [16] constrains the procedure by searching for the apices that best fit a priori information on the factors when they are available. Benali *et al.* [15] introduces also a statistical model to achieve an optimal separation between signal and noise for scintigraphic dynamic studies.

Nijran and Barber [8] [9] use an intersection method. Besides the study space that represents the studied trixels, they define a theory space generated from a theoretical model of all the possible trixels that can describe the physiological structure of interest. The factor of interest lies in both the spaces simultaneously. However, this method is strongly dependent on the relevance of the theoretical model of the structure of interest.

Some authors [10] [33] apply a method based on the maximum entropy principle.

Benali *et al.* [19] unifies some previous approaches ([16] [15]) and the hypothesis related to the fundamental factor curves and images in scintigraphic dynamic studies.

Some studies [36] [55] are based on specific constraints that reflect prior knowledge on their studies and thus, are only adapted to their studies.

Martel *et al.* [41] [42] incorporate additional constraints on factor images and factor curves appropriate for DCE-MRI studies

Some groups [45] [46] [58] use a penalized least squares objective function to minimize the non negative terms responsible for the overlapping between factor volumes.

In the study by Martel *et al.* [49], the factor curves are determined by the operator who selects representative TIC of anatomic structures.

4.7.3 Proposed physiological constraints on the factors

The constraints which are proposed should be tight enough to prevent from ambiguous solution but sufficiently general to ensure that all kinds of MRU studies can be considered, even those where pathological structures enhance.

So, the following constraints are incorporated and do not restrict the algorithm to a specific type of study:

- **Positivity constraints on the factor volumes.** Considering TIC, the contribution of each factor coefficient on factor curves can only be positive: $\forall n=1\dots N$ and $\forall k=1\dots K$, $a_{nk}\geqslant 0$.
- **Remove ambiguity on scale factor.** Without any constraint on $\boldsymbol{R}$, the factors $\boldsymbol{A}$ and $\boldsymbol{F}$ are determined up to a scale factor. So, the following constraints are added to keep the true magnitude of the factor curves:
  - □ $\forall n=1\dots N$, $\sum_{k=1}^{K} a_{nk}=1$
  - □ $\forall k=1\dots K$ and $\forall t=1\dots T$, $\varphi_t\leqslant f_{kt}\leqslant\phi_t$

    φ_t and ϕ_t are determined by averaging respectively 1 % of the lowest and 1 % of the greatest quadrixels at time t.
  - □ So, ideally, neglecting the error term, a pure kidney quadrixel $\boldsymbol{d}_{n.}$ would be expressed from (4.7.2) as $\boldsymbol{d}_{n.}=\boldsymbol{f}_{1.}+\boldsymbol{\varepsilon}_{n.}$ with $a_{n1}=1$, $a_{n2}=0$ and $a_{n3}=0$. If TIC are considered, this constraint is more restricting than a simple positivity constraint on $\boldsymbol{F}$ because usually $\varphi_t>0$. But if TEC were considered, the constraint would take into account that f_{kt} can also be negative.
- **Reduce overlapping.** Ideally, each factor volume should describe a single physiological structure. But due to the non unique solution of FAMIS (Sitek *et al.* [46] explains nonuniqueness of the factor model), a factor volume may be a linear combination of a number of true physiological volumes. To reduce the amount of mixing between factors, all quadrixels are forced to be mainly explained by a single set of factor volume and factor curve by keeping to a high value the greatest factor coefficient and keeping to a low value the others.

Our proposed iterative procedure to solve for $\boldsymbol{A}$ and $\boldsymbol{F}$ can now be stated.

4.7.4 Proposed FAMIS algorithm

The steps are summarized in Figure 22 and are as follows:

1. Compute $\boldsymbol{\varphi}=\left[\varphi_1 \quad \dots \quad \varphi_T\right]$, $\boldsymbol{\phi}=\left[\phi_1 \quad \dots \quad \phi_T\right]$ and $\boldsymbol{V}$, $\boldsymbol{\Lambda}$, $\boldsymbol{U}$ from (4.7.5).
2. Compute a first estimate of $\boldsymbol{A}$ from $\boldsymbol{A}=\boldsymbol{V}\ \boldsymbol{\Lambda}$.
3. Adjust $\boldsymbol{A}$ to satisfy the constraints:
   - Replace all negative elements of $\boldsymbol{A}$ by 0.01.
   - Normalize each row of $\boldsymbol{A}$ so that $\forall n=1\dots N$, $\sum_{k=1}^{K} a_{nk}=1$.
   - For each row $n=1\dots N$ of $\boldsymbol{A}$,
     - if the K factor coefficients a_{nk} are smaller than 0.5 then, set $a_{nk^*}=1-(0.01\cdot K)$ where a_{nk^*} is the greatest of the K factor coefficients and set $a_{nk}=0.01$ to the other factor coefficients $k\neq k^*$.
     - next n
   - Normalize again each row of $\boldsymbol{A}$ so that $\forall n=1\dots N$, $\sum_{k=1}^{K} a_{nk}=1$.
4. From the adjusted $\boldsymbol{A}$, compute $\boldsymbol{R}$ given by (4.7.6) as $\boldsymbol{R}=(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'(\boldsymbol{V}\ \boldsymbol{\Lambda})$ and hence, compute $\boldsymbol{F}$ given by (4.7.7) as $\boldsymbol{F}=\boldsymbol{R}\ \boldsymbol{U}$.
5. Adjust $\boldsymbol{F}$ to satisfy the constraints:
   - $\forall k=1\dots K$ and $\forall t=1\dots T$, if $f_{kt}<\varphi_t$ then set the value of f_{kt} to φ_t and if $f_{kt}>\phi_t$ then set the value of f_{kt} to ϕ_t.
6. From the adjusted $\boldsymbol{F}$, compute $\boldsymbol{R}^{-1}$ given by (4.7.7) as $\boldsymbol{R}^{-1}=\boldsymbol{U}\boldsymbol{F}'(\boldsymbol{F}\boldsymbol{F}')^{-1}$ and hence, compute $\boldsymbol{A}$ given by (4.7.6) as $\boldsymbol{A}=\boldsymbol{V}\ \boldsymbol{\Lambda}\ \boldsymbol{R}^{-1}$.
7. If the change between two consecutive factor volumes $\boldsymbol{A}$ or factor curves $\boldsymbol{F}$ is less than a given threshold, then stop the procedure. Else return to 3.

It shall be noted that even at the end of the procedure, $\boldsymbol{A}$ and $\boldsymbol{F}$ may not both simultaneously satisfy all the physiological constraints because of noise and outliers. The final values should be regarded as a trade-off between the constraints on $\boldsymbol{A}$ and the constraints on $\boldsymbol{F}$.

Moreover, the original data set $\boldsymbol{D}$ can be reconstructed, with an error term that can be attributed to noise, from the factor volumes $\boldsymbol{A}$ and the factors curves $\boldsymbol{F}$ (4.7.4) and thus, makes also FAMIS a useful method for compression.

The factor volumes $\boldsymbol{A}$ can be interpreted as a factor coefficient map. The greater value a_{nk}, the more quadrixel $\boldsymbol{d}_{n.}$ is made of factor curve $\boldsymbol{f}_{k.}$. Like fuzzy K-means and EM clustering, an operator can select a tissue membership threshold to obtain, from the factor coefficients map, a hard segmentation of the physiological structures.

Finally, our proposed FAMIS method does not need any random initialization and consequently, the results are reproducible whereas all clustering methods are dependent on the selection of starting values.

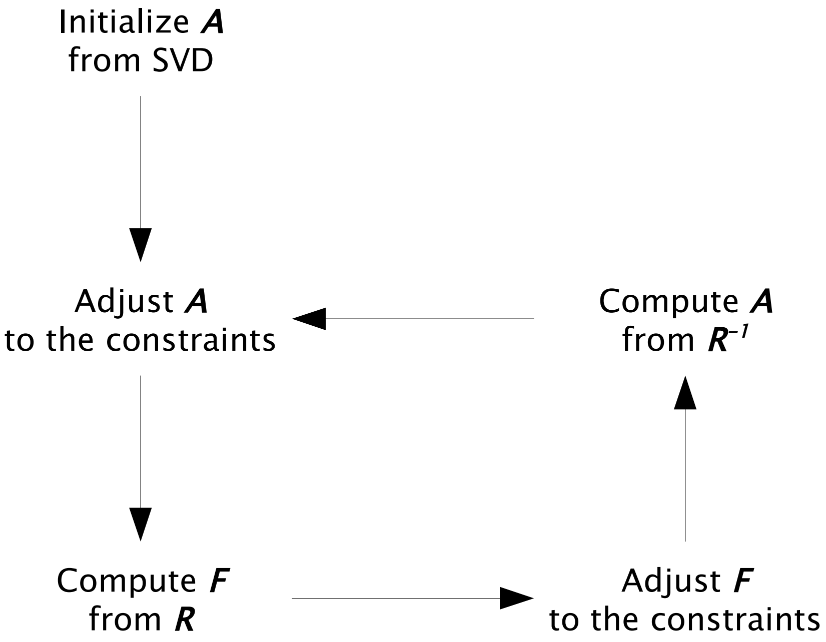


Figure 22: FAMIS algorithm

If the change between two consecutive factor volumes A or factor curves F is less than a given threshold, then the procedure is stopped.

4.8 Summary

4.8.1 Common points between our approaches

Although clustering analysis (KNN, FKM, EM) and factor analysis (FAMIS) are two different kinds of approaches for kidney segmentation, they share surprisingly some common points:

- Fundamental factor curves in FAMIS can be considered as the equivalent of the representative TIC or weights in clustering methods: a factor curve describes the main kinetic in a physiological region while the weights represent the average TIC in this same region.
- Factor volumes in FAMIS can be considered as the equivalent of the clusters in clustering method as they both correspond in the dynamic image sequence to physiological regions sharing similar kinetics.
- Like fuzzy memberships in fuzzy K-means clustering and probabilities in EM clustering, factor volumes in FAMIS give a soft partition of the physiological regions.
- FAMIS and EM clustering both assume a linear decomposition of a limited number of physiological regions (observe the similarity of equations 4.6.2 and 4.7.1).

4.8.2 Summary of our proposed methods

A summary of our 4 proposed methods to segment kidney tissues is given by Figure 23.

At the end, in an automatic step, the volumetric structures of interest can finally be segmented by extracting only the biggest 6-connected objects. But an extra step is necessary in rare cases where the kidney is so severely damaged that some sections do not link into a coherent object.

Method	Approach	Model / Objective function	Computation	Output
Kohonen Neural Network	Non-parametric clustering		Based on Euclidean distances	Hard clustering: data points are assigned to the nearest cluster centre
Fuzzy K-Means clustering	Non-parametric clustering	Minimize the sum of the squared distances weighted by membership functions (4.5.1)	Based on Euclidean distances weighted by membership functions	Fuzzy membership map: a higher membership value gives more confidence
EM clustering	Parametric clustering (a Gaussian mixture density is assumed)	Best fit the distribution of the data set by a Gaussian mixture density (4.6.5)	Based on probabilities	Probability map: a higher probability gives more likelihood
FAMIS	Factor analysis	Data set is a linear combination of factor curves and factor volumes (4.7.4)	Based on SVD and physiological constraints	Factor coefficient map: a higher coefficient gives more contribution

Figure 23: Summary of proposed methods

5 Experimental results and Discussion

5.1 Introduction

Our proposed methods are performed on 20 real MR data sets that were all collected between 2003 and 2006 at the National University Hospital. The 20 data sets represent a wide range of possible scenarios, from normal to very poorly functioning kidneys (refer to **Appendix** for a brief description of the data sets).

The data sets consist of 24 to 44 volumes, each volume containing 10 to 18 slices whose image matrix size is 256x256.

All our methods are coded in MATLAB(R) 7.0, The MathWorks, Inc. Kohonen neural network (KNN) and Fuzzy K-means clustering (FKM) are performed using the Matlab toolboxes (respectively neural network toolbox 4.0 and fuzzy logic toolbox 2.0) with the default values (that is, a learning rate $\alpha=0.01$ for KNN and a weighting exponent $m=2$ for FKM) while Expectation-Maximization clustering (EM) and FAMIS are implemented on our own.

5.2 Parameter settings

5.2.1 Initialization

Background elimination

No drawing region is required from the user except for data set 041124pa3 where a ROI is manually drawn to exclude the aorta and the right femoral artery.

For the three clustering techniques (KNN, FKM, EM), the background is eliminated by excluding quadrixels whose mean TEC over time is below an enhancement threshold $\Theta_E=0.5$. The value is chosen low enough so that it does not exclude any kidney voxels, but it is also high enough to exclude from analysis most of the irrelevant tissues (muscle, guts, bones, ...) that do not enhance.

For FAMIS, the strategy is different. This technique has not the ability to separate the left kidney cortex from the spleen in most data sets. So, there is an attempt to eliminate the splenic region beforehand by using a high enhancement threshold value. But, a too high value is not advisable since it would discard from analysis renal voxels too. That is why an enhancement threshold Θ_E on TEC is not applied but rather an intensity threshold Θ_I on TIC: in terms of relative contrast enhancement, the spleen signal intensity is high and consequently, a too high enhancement threshold value Θ_E on TEC would be required to exclude the splenic region. Hence, before submitting the data sets to FAMIS analysis, quadrixels whose mean TIC over time is below an intensity threshold Θ_I are discarded. Using a threshold on mean TIC rather than TEC is worthwhile for FAMIS because it gives the possibility to exclude the spleen region from analysis. But it has the disadvantage that it requires the user to select a suitable threshold value Θ_I for each data set: Θ_I must be chosen high enough to discard splenic voxels, but it must be low enough to keep renal voxels for analysis. To aid the user in selecting Θ_I, a simple graphical user interface (GUI) is used to adjust the threshold level and preview the result for one particular slice.

Number of clusters and number of factors

The number of factors in FAMIS is chosen to be $K=5$ while the number of clusters in clustering methods is chosen to be $K=6$. The same number $K=5$ for clustering methods could be used as well, but seeking for one more cluster is considered to be more appropriate as a low threshold value $\Theta_E=0.5$ is used: the additional cluster can regroup the noise or other irrelevant tissues.

In case of doubt, it is usually more advisable to overestimate the number of clusters or factors than to underestimate it. Indeed, if the number is underestimated, there is a risk that several tissue types are regrouped into one single cluster or factor and thus, separating the different tissue types would require to split up the cluster or factor, which is a difficult task. On contrary, if it is overestimated, one tissue type may certainly be regrouped into several clusters or factors, but it would be an easy task to merge them.

Figure 24 summarizes the different parameter values to initialize the clustering techniques (KNN, FKM, EM) and FAMIS for each of the 20 data sets.

	Clustering techniques		FAMIS	
Data sets	Threshold on TEC	Number of clusters	Theshold on TIC	Number of factors
031104pa5	0.5	6	130	5
031113e1pa1			35	
031223pa3			50	
040205bpa1			45	
040205pa1			110	
040205pa2			70	
040426apa1			80	
040426dpa1			80	
040427pa1			45	
040430pa1			105	
040831cpa1			40	
040902pa1			30	
041126pa2			60	
041126pa3*			60	
041126pa4			35	
050518epa2			20	
050613pa2			35	
060721apa3			75	
060721bpa4			45	
060721cpa1			75	

Figure 24: Initialization parameters for clustering techniques and FAMIS
Only data set 041126pa3 requires a region of interest (ROI) to be drawn to exclude aorta and right femoral artery.*

5.2.2 Volume extraction

KNN provides a hard segmentation: it outputs $K=6$ binary volumes and the interaction of an operator is just needed to associate each volume to a possible tissue type based on visual inspection. On contrary, FKM, EM and FAMIS provide a soft segmentation: FKM outputs fuzzy membership values, EM outputs probabilities and FAMIS outputs factor coefficients. Although clinical works may require a volume to be the sum of voxels with partial tissue membership, for the purposes of algorithm comparison, all the output memberships are rendered to be hard. Hence, for FKM, EM and FAMIS, the output must be “defuzzified” to enforce a binary decision. That is, for FKM and EM, the voxels must be assigned to one single cluster while for FAMIS, they must be assigned to one single factor to obtain a final hard segmentation. This decision could be done automatically by assigning each voxel uniquely to the cluster or factor where it has the highest membership value (or probability or factor coefficient). But in the proposed approach, an operator is required to threshold by visual inspection the tissue membership maps (or probability maps or factor coefficient maps) to yield binary maps so that the uncertain information can be dealt with more flexibility. The tissue membership threshold operation is applied individually to the right and left kidneys and the values used for each data set are given in Figure 25.

	FKM		EM		FAMIS		FAMIS	
Data sets	Left Parenchyma	Right Parenchyma	Left Parenchyma	Right Parenchyma	Left Parenchyma	Right Parenchyma	Left Cortex	Right Cortex
031104pa5	0.2	0.5	0.8	0.4	0.1	0.1	0.3	0.3
031113e1pa		0.3		0.3		0.9		0.7
031223pa3	0.1	0.8	0.1	0.6	0.9	0.8	0.8	0.8
040205bpa1	0.4	0.4	0.6	0.8	0.5	0.5	0	0.4
040205pa1	0.2	0.4	0.4	0.6	0.7	0.8	0.6	0.7
040205pa2	0.2	0.2	0.1	0.6	0.1	0.5	0.3	0.5
040426apa1	0.6	0.1	0.7	0.3	0.1	0	0.3	0
040426dpa1	0.3	0.2	0.2	0.1	0.5	0.1	0.4	0.1
040427pa1	0.2	0.2	0.1	0.1	0.2	0.6	0.1	0.5
040430pa1	0.2	0.4	0.1	0.5	0.1	0.8	0.3	0.7
040831cpa1	0.1	0.1	0.3	0.2	0.5	0.5	0.2	0.4
040902pa1	0.2	0.1	0.9	0.6	0.9	0.6	0.6	0.5
041126pa2	0.3	0.2	0.2	0.1	0.5	0.4		
041126pa3	0.4	0.1	0.8	0.2	0.7	0.2		
041126pa4	0.2	0.2	0.5	0.7	0.4	0.9	0.1	0.2
050518epa2	0.1	0.2	0.4	0.1	0.6	1.5	0.6	0.7
050613pa2	0.1	0.2	0.1	0.2	0.8	0.8	0	0.4
060721apa3	0.2	0.1	0.1	0.1	0.4	0.4	0.2	0.3
060721bpa4	0.1	0.7	0.1	0.3	0.1	0.5	0.1	0.5

Figure 25: Tissue membership threshold values used to “defuzzify” renal parenchyma and cortex volumes

- *Data set 031113e1pa1 has no left kidney, so no left parenchyma and no left cortex are extracted.*
- *FAMIS could not extract from data sets 041126pa2 and 041126pa3 any renal cortex.*

Finally, the interested volumes that contain renal tissues are extracted from the binary maps by considering the biggest 6-connected component over the slices. For most of the data sets, the volume extraction does not include the most anterior and posterior slices because they do not contain renal regions but rather some irrelevant tissues like bowels.

An example of tissue extraction for left and right renal parenchyma from EM and FAMIS is given in Figure 26 and Figure 27. Among all the volumes output by our methods, the one associated to the parenchyma must be identified. Merging a number of volumes to obtain the parenchyma volume is also possible, especially for our FAMIS method which has the ability to return the parenchyma in two different factor volumes (cortex and medulla factor volumes). Then, a tissue membership threshold is applied and the biggest 6-connected component over the slices is extracted to yield the final parenchyma segmentation.

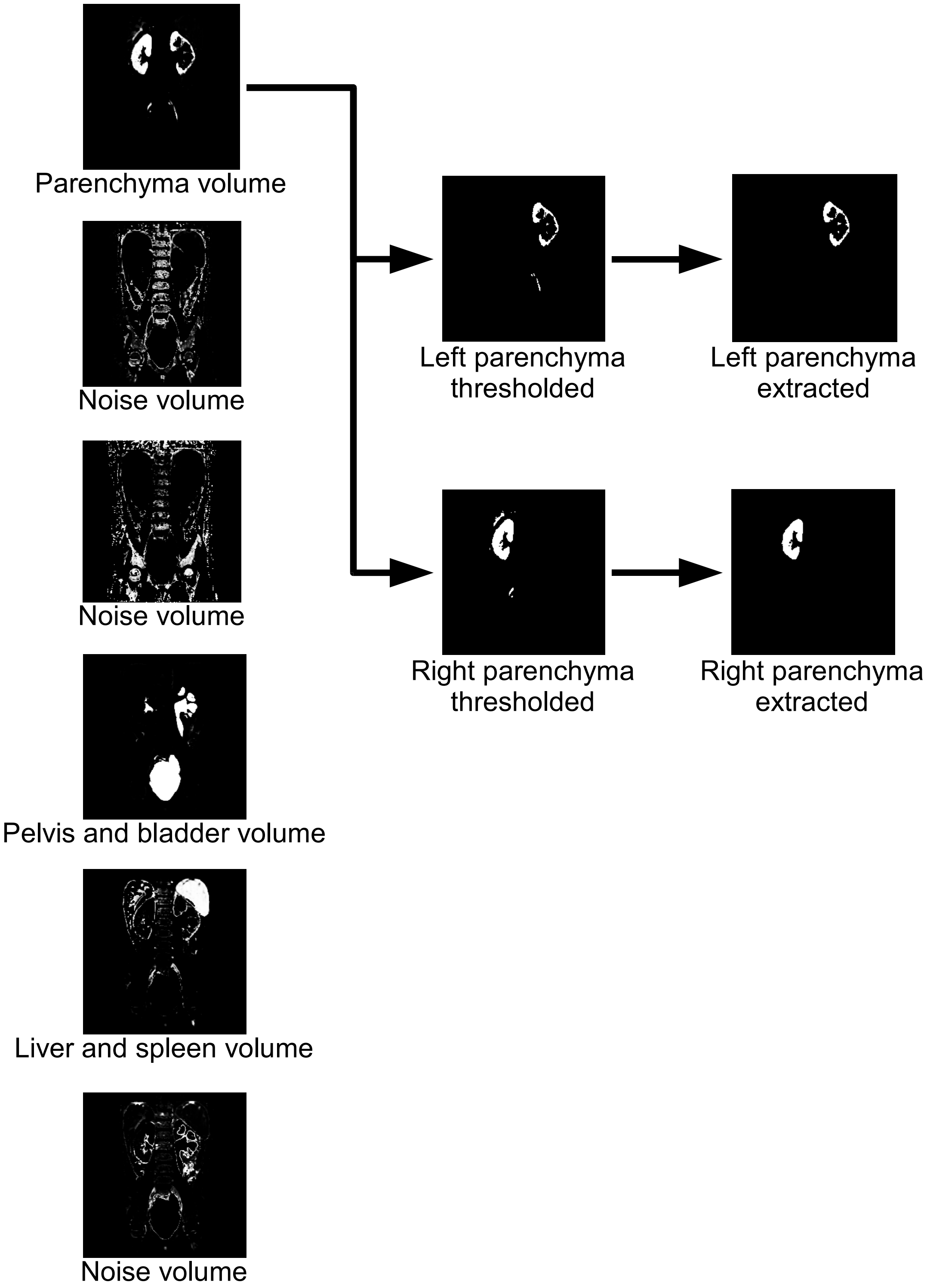


Figure 26: Extraction of renal parenchyma volume in EM

- *Each of the 6 volumes output by EM may be associated to a possible anatomical tissue and the one associated to the renal parenchyma is identified.*
- *The parenchyma volume is described by probabilities that reflect tissue memberships. A tissue membership threshold is applied to obtain a hard segmentation. The thresholding operation is applied individually to left and right parenchyma.*
- *Finally, left and right parenchyma volumes are obtained by extracting the biggest 6-connected component over the slices.*
- *Data set 060721cpa1 (only one single slice of the volumes are shown).*

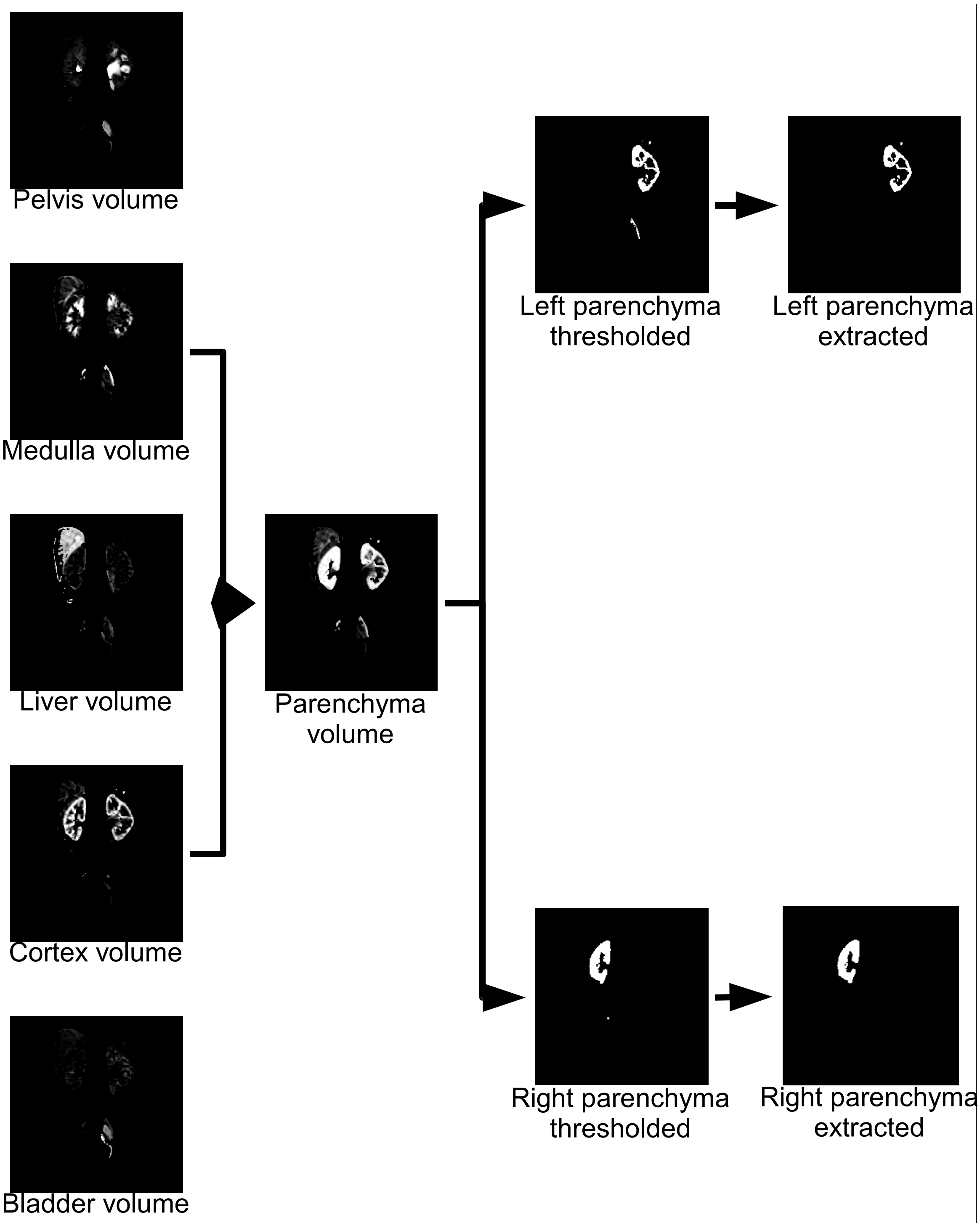


Figure 27: Extraction of renal parenchyma volume from factor volumes in FAMIS

- *Each of the 5 factor volumes output by FAMIS may be associated to a possible anatomical tissue. The ones associated to the renal cortex volume and medulla volume are identified and are mathematically added to yield a renal parenchyma volume.*
- *The parenchyma volume is described by factor coefficients that reflect tissue memberships. A tissue membership threshold is applied to obtain a hard segmentation. The thresholding operation is applied individually to left and right parenchyma.*
- *Finally, left and right parenchyma volumes are obtained by extracting the biggest 6-connected component over the slices. Left and right cortex volumes could be obtained in the same way.*
- *Data set 060721cpa1 (only one single slice of the volumes are shown).*

5.3 Result examples

5.3.1 Data set 031113e1pa1

Data set 031113e1pa1 has no left kidney while the right one has a very particular shape. A segmentation approach guided by a prior knowledge of a typical kidney shape is unlikely to work for this data set due to the very unique shape of the right kidney. On contrary, our proposed methods achieve good results (Figure 28).

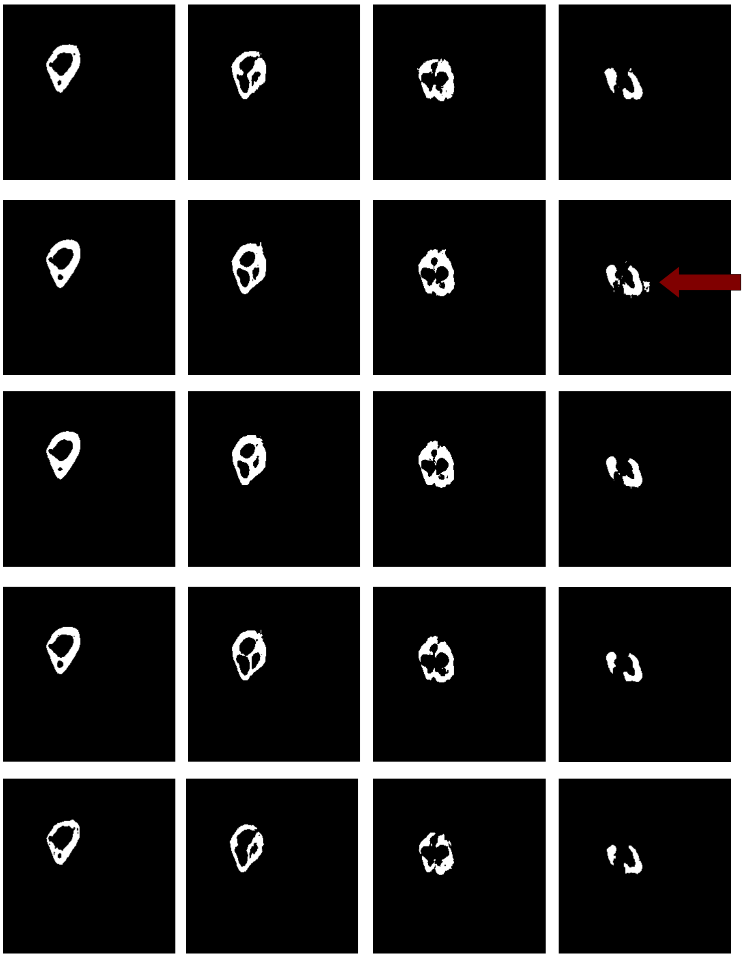


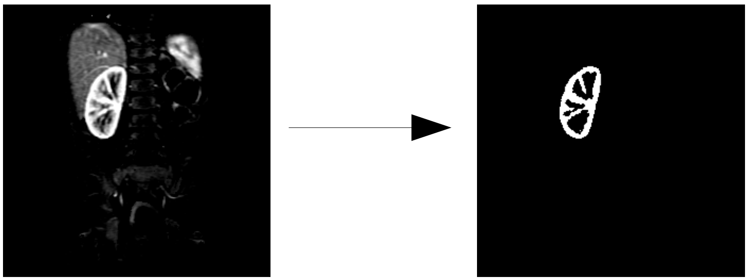
Figure 28: Results for data set 031113e1pa1

- *From 1st to 5th row: ground truth, KNN, FKM, EM and FAMIS results.*
- *From 1st to 4th column: slice 8, 10, 12 and 14.*
- *The data set is characterised by the absence of left kidney and the particular shape of right kidney. Yet, our techniques based on temporal signal intensities achieve good results.*
- *Note that part of the aorta (arrow) is included in the KNN result on slice 14. But it is removed in other techniques by selecting a proper tissue membership threshold.*

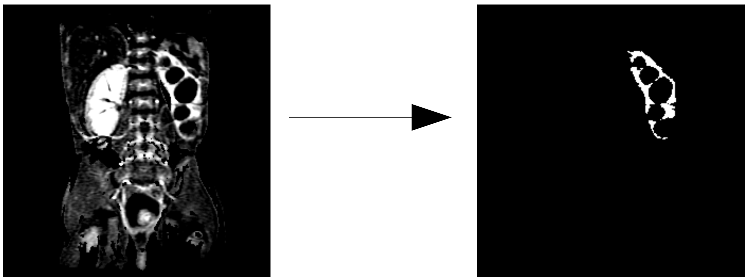
5.3.2 Data set 050518epa2

If both right and left kidneys were functioning similarly, they would have similar signal intensities, and thus, they would be expected to be regrouped in the same cluster or factor volume. But, data set 050518epa2 is a case where only the right kidney is functioning normally while the left one is functioning poorly and has a strange shape (which may be signs of hydronephrosis). Hence, they have very different signal intensities and thus, our proposed methods separate them in different clusters or different factor volumes.

Figure 29 shows two of the five factor volumes output by FAMIS. The first factor volume can be mainly identified to the right kidney cortex (and the spleen). However, the left kidney cortex is not present in the same factor volume because it enhances abnormally over time. Indeed, it enhances later and thus, it appears more like a normal renal medulla. That is why the left kidney cortex is extracted from the factor volume that contains the right renal medulla.



Factor Volume 1 Right renal cortex extracted



Factor Volume 2 Left renal cortex extracted

Figure 29: FAMIS results for data set 050518epa2

Right renal cortex and left renal cortex are extracted from 2 different factor volumes, which reflects their difference in physiology (the right renal cortex is functioning normally whereas the left renal cortex is malfunctioning and has a strange shape).

5.3.3 Data set 041126pa3

Data set 041126pa3 is the example of a patient whose both kidneys are very poorly functioning: the larger right kidney appears to have a slower transit than the smaller left kidney. No functioning renal medulla is visible and no excretory phase occurs over the acquisition time of 22 minutes, revealing probably severe kidney obstruction. Most of the methods based on thresholding or deformable model are unlikely to give satisfying results for those two kidneys. This is the only data set that is handled by requiring from the operator to draw a region of interest that excludes the aorta and the right femoral artery (Figure 30).

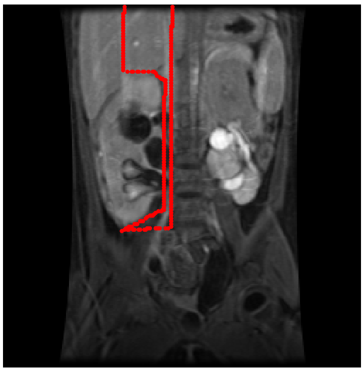


Figure 30: Drawing ROI for data set 041126pa3

- *An operator is required to draw a ROI on a maximum intensity projection (MIP) of slice 8 derived from the dynamic sequence.*
- *The volume inside the drawing ROI, which includes aorta and right femoral artery, is excluded from analysis for all our methods.*

By disconnecting the aorta and the right femoral artery from the right kidney, our proposed methods can achieve results reasonably close to the ground truth. Especially, FAMIS yield factor curves that may provide useful information to the clinician (Figure 31): the kidney factor curve, which is supposed to be upper concave in normal pattern, is continually rising and flattened because of a possible persisting tubular phase. Consequently, those patient kidneys may be suspicious of very severe obstruction and may draw the attention of the clinician.

It must be noted that for this data set, both the kidneys are functioning so differently from their normal pattern that all our methods, including FAMIS, can easily differentiate the splenic tissue from the renal tissue.

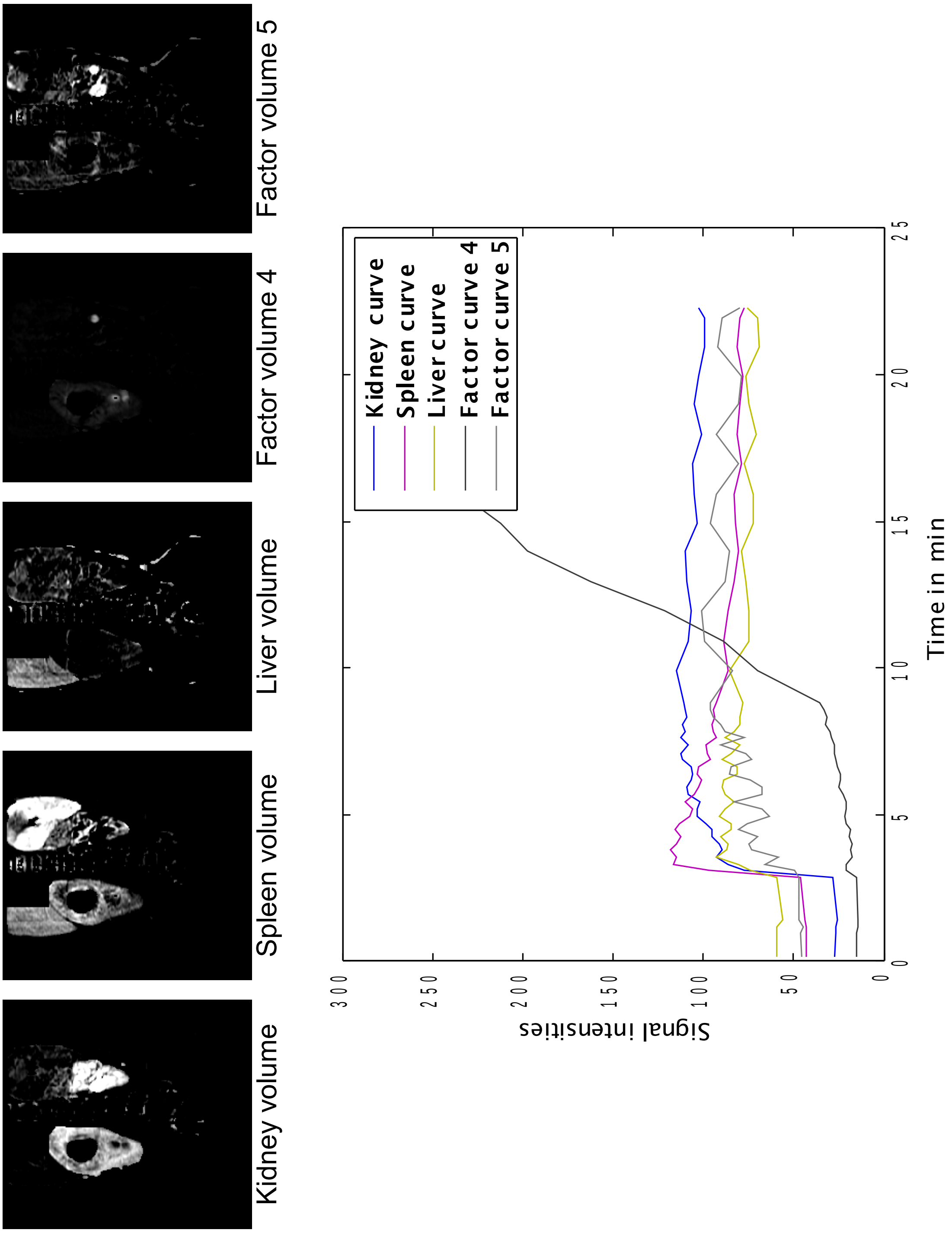


Figure 31: Factor volumes and factor curves for data set 041126pa3

- *FAMIS returns a reasonably good kidney factor volume. But the aorta and the right femoral artery had to be removed from analysis by an operator interaction, otherwise they would still be connected to the right kidney and thus, would yield poor segmentation results.*
- *Examination of factor volumes and factor curves draw attention on the abnormal kidney pattern, possibly caused by severe obstruction .*

5.3.4 Data sets 040427pa1 and 040902pa1

Unfortunately, there are data sets where the clustering methods achieve less satisfactory kidney segmentation compared to the ground truth. Data set 040427pa1 and 040902pa1 are two examples of those data sets. The main cause of failure is due to the presence of the spleen or the liver with respectively the left or right kidney (Figure 32). This misclassification happens when the kidney tissue has similar signal intensities with the surrounding organs such as spleen and liver, and consequently, our clustering methods can hardly regrouped them in separate clusters. This is more likely to happen for poorly functioning kidneys. Yet, our FAMIS method could return reasonably good results for all our 20 MR data sets.

KNN result
on data set 040427pa1

FAMIS result
on data set 040427pa1

Ground truth
on data set 040427pa1

EM result
on data set 040902pa1

FAMIS result
on data set 040902pa1

Ground truth
on data set 040902pa1

Figure 32: Examples of segmentation failures by clustering methods

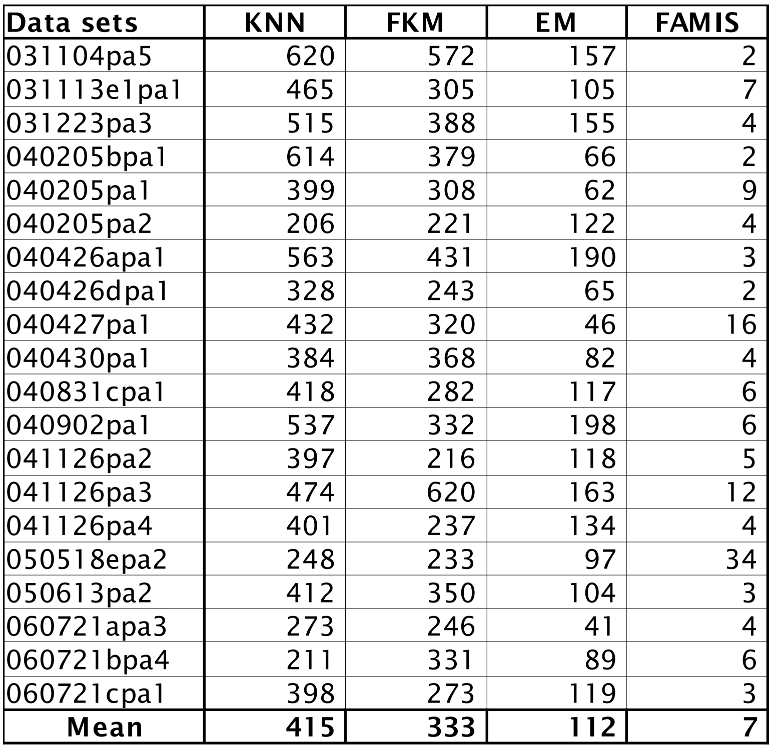
- *Ground truth is the manual segmentation performed by the expert.*
- *The segmentation result on data set 040427pa1 (slice 5) by KNN contains both right parenchyma and part of the liver tissue: KNN fails to regroup right parenchyma and liver tissues in separate clusters based on their signal intensities. On the contrary, FAMIS achieves a reasonably good result for this data set.*
- *The segmentation result on data set 040902pa1 (slice 7) by EM contains both left parenchyma and part of the splenic tissue: EM fails to regroup left parenchyma and splenic tissues in separate clusters based on their signal intensities. On the contrary, FAMIS achieves a reasonably good result for this data set.*

5.4 Performance evaluation

5.4.1 Quantitative measures

Computation time

Our proposed methods are run in Matlab on a Intel Pentium(R) D 2.8GHz with 1Gb Ram. The computation time needed to process each of the 20 data sets are given in Figure 33.



Data sets	KNN	FKM	EM	FAMIS
031104pa5	620	572	157	2
031113e1pa1	465	305	105	7
031223pa3	515	388	155	4
040205bpa1	614	379	66	2
040205pa1	399	308	62	9
040205pa2	206	221	122	4
040426apa1	563	431	190	3
040426dpa1	328	243	65	2
040427pa1	432	320	46	16
040430pa1	384	368	82	4
040831cpa1	418	282	117	6
040902pa1	537	332	198	6
041126pa2	397	216	118	5
041126pa3	474	620	163	12
041126pa4	401	237	134	4
050518epa2	248	233	97	34
050613pa2	412	350	104	3
060721apa3	273	246	41	4
060721bpa4	211	331	89	6
060721cpa1	398	273	119	3
Mean	**415**	**333**	**112**	**7**

Figure 33: Computation time in seconds

It is clear that FAMIS has the best computation time efficiency: it returns the factor volumes and the factor curves within seconds. This can be explained by the simple matrix computations involved in the algorithm and by the high threshold in preprocessing that excludes from analysis a great amount of data, essentially the spleen and the noise.

KNN is the most computationally expensive method because it is a single sample update algorithm. That is, the weights are updated after each quadrixel is presented to the neural network. On contrary, FKM and EM are faster because they are batch algorithms: they wait until all quadrixels have been re-clustered before they update their cluster centres.

Finally, EM processes faster than FKM mainly because its data is preprocessed by a PCA to reduce the computation complexity of the algorithm.

Segmentation results compared to ground truth

The proposed methods are validated by using a measure that quantifies the discrepancy between our segmentation results and the tracings of cortex and whole parenchyma by a human expert (Dr Borys Shuter from the National University Hospital). This is done by computing the undirected average distance:

$$H(A,B)=\max(h(A,B),h(B,A)) \tag{5.4.1}$$

$$h(A,B)=\frac{1}{N_A}\sum_{a\in A}\min_{b\in B}\|a-b\| \tag{5.4.2}$$

- A is the manual segmentation, called ground truth, performed by the expert.
- N_A is the number of voxels in A.
- B is the segmentation result obtained from our proposed methods.

The evaluation results for each of the 20 MR data sets are tabulated in Figure 34.

The segmentation performance of each of our proposed methods can be compared to each other to determine whether there is one method that achieves closer results to the ground truth than another one: a Student *t*-test on paired two samples is used with the null hypothesis H_0 that $\mu_1=\mu_2$ against the alternate hypothesis H_1 that $\mu_1>\mu_2$, where μ_1 is the mean of the undirected average distances achieved by one of our proposed methods and μ_2 is that by another one.

Two comparative tests are statistically significant: it is found that $T_{statistic}>t_{critical}$ when KNN is compared to FKM and when KNN is compared to FAMIS (Figure 35). Hence, for those two cases, H_0 is rejected with a confidence of 95%. That is, KNN results are significantly more different from the ground truth than FKM and FAMIS results.

On contrary, comparative tests between our other proposed methods do not give us enough evidence to reject H_0.

	KNN		FKM		EM		FAMIS		FAMIS	
Data sets	Left Parenchyma	Right Parenchyma	Left Parenchyma	Right Parenchyma	Left Parenchyma	Right Parenchyma	Left Parenchyma	Right Parenchyma	Left Cortex	Right Cortex
031104pa5	1.29	1.28	1.27	1.09	1.24	1.14	1.29	1.3	0.88	0.79
031113e1pa1		0.91		0.74		0.79		1.13		0.9
031223pa3	**2.11**	0.84	**2.19**	0.6	2	0.55	1.85	0.92	1.68	0.93
040205bpa1	0.88	1.29	1.14	1.48	1.17	1.57	1.11	1.41	1.35	0.83
040205pa1	0.65	1.27	0.63	1.02	0.66	1.19	0.57	0.87	0.48	0.6
040205pa2	1.65	1.36	1.36	1.12	1.08	1.29	0.76	1.01	0.89	0.81
040426apa1	1.16	1.21	1.08	0.85	1.02	1.25	0.76	1.37	0.74	0.68
040426dpa1	0.6	1.05	0.55	0.98	0.6	0.94	0.61	1.06	0.61	1.11
040427pa1	**2.06**	**3.56**	1.49	1.32	1.5	1.4	1.05	1.05	0.9	0.52
040430pa1	0.82	1.73	0.81	1.34	0.82	1.12	0.75	0.96	0.6	0.68
040831cpa1	1.18	0.7	1.21	0.67	1.3	0.99	0.95	0.76	1.05	1.1
040902pa1	1.47	1.59	1.04	1.63	**5.96**	1.77	1.13	1.23	1.83	1.23
041126pa2	0.84	0.97	0.89	0.95	0.83	0.96	0.85	1.05		
041126pa3	0.88	1.71	0.75	1.51	1.08	1.36	1.4	1.01		
041126pa4	1.19	0.84	1.16	0.82	1.29	1.02	0.87	1.32	0.84	1.04
050518epa2	**3.72**	1.02	**3.89**	1.06	**3.27**	0.76	1.72	0.78	1.72	0.64
050613pa2	1.36	0.78	1.37	0.58	1.41	0.69	1.43	0.65	1.35	0.93
060721apa3	1.68	1.41	1.66	1.25	1.66	1.37	1.44	1.16	1.62	0.83
060721bpa4	**2.21**	1.3	1.48	0.78	0.94	1.05	1.11	0.97	1.76	1.22
060721cpa1	0.92	0.71	0.83	0.59	0.8	0.61	0.98	1.25	1.26	0.76
Mean	**1.4**	**1.28**	**1.31**	**1.02**	**1.51**	**1.09**	**1.09**	**1.06**	**1.15**	**0.87**

Figure 34: Evaluation results using undirected average distance (unit: voxel)

- *The renal parenchyma (left and right) obtained from KNN, FKM, EM and FAMIS for each data set are compared to the parenchyma ground truth using the undirected average distance.*
- *The renal cortex (left and right) obtained from FAMIS for each data set are also compared to the cortex ground truth using the undirected average distance.*
- *Data set 031113e1pa1 has no left kidney, so no left parenchyma and no left cortex are extracted.*
- *FAMIS could not extract from data sets 041126pa2 and 041126pa3 any renal cortex.*

	KNN vs FKM	KNN vs FAMIS
$T_{statistic}$	2.83	2.73
p value	0.005	0.004

Figure 35: Results of the Student *t*-tests for renal parenchyma

5.4.2 Qualitative measure

FAMIS has the strong advantage over all the clustering methods that it is the only method which has the ability to provide separately the renal cortex and the renal medulla. Among the 20 MR data sets, there are only 2 cases where FAMIS cannot distinguish the cortical and medullary tissues but still provides the whole renal parenchyma. In the first case (data set 041126pa3), kidney cortex and medulla are severely compromised and no clear signal differences can be measured. In the second case (data set 041126pa2), the reasons are still unknown, but it might be due to a too slow contrast agent injection to the patient.

To give another qualitative measure of the advantage of FAMIS over all the clustering methods, it can be emphasized that FAMIS yields factor curves that are physiologically interpretable by the clinician. Indeed, Figure 36 shows that FAMIS can extract physiological compartments and their associated physiological factor curves which clearly reveal the different phases of the renal enhancement pattern:

- A pre-enhancement phase where the liver tissue is hyperintense for about the first 3 minutes.
- A vascular phase characterised by a cortex intensity peak and followed, after a slight delay, by an intensity peak in the liver.
- A subsequent tubular phase where the signal intensity of the renal medulla exceeds that of the cortex to reach a peak corresponding to the arrival of contrast agent in the loop of Henle, and then decreases homogeneously with the cortex.
- An excretory phase where the renal pelvis and bladder tissues enhance.

On contrary, all the clustering techniques extract the entire renal parenchyma without differentiating the renal cortex and renal medulla. Moreover, the TIC which are associated to the cluster centres and are representative of the compartments are more difficult to interpret physiologically. Figure 37 shows that clustering methods extract:

- A parenchyma curve, which is a mixture of both cortical and medullary signals, that reaches the highest amplitude.
- A pelvis and/or bladder curve.
- A splenic curve whose signal intensity is lower than that of the parenchyma.

- 2 or 3 other curves that are only different from each other in amplitude. They do not describe any particular physiological compartments and may rather correspond to the noise present in the dynamic sequence.

In that sense, clustering methods tend to regroup the quadrixels based more on their signal amplitude rather than their characteristic physiology.

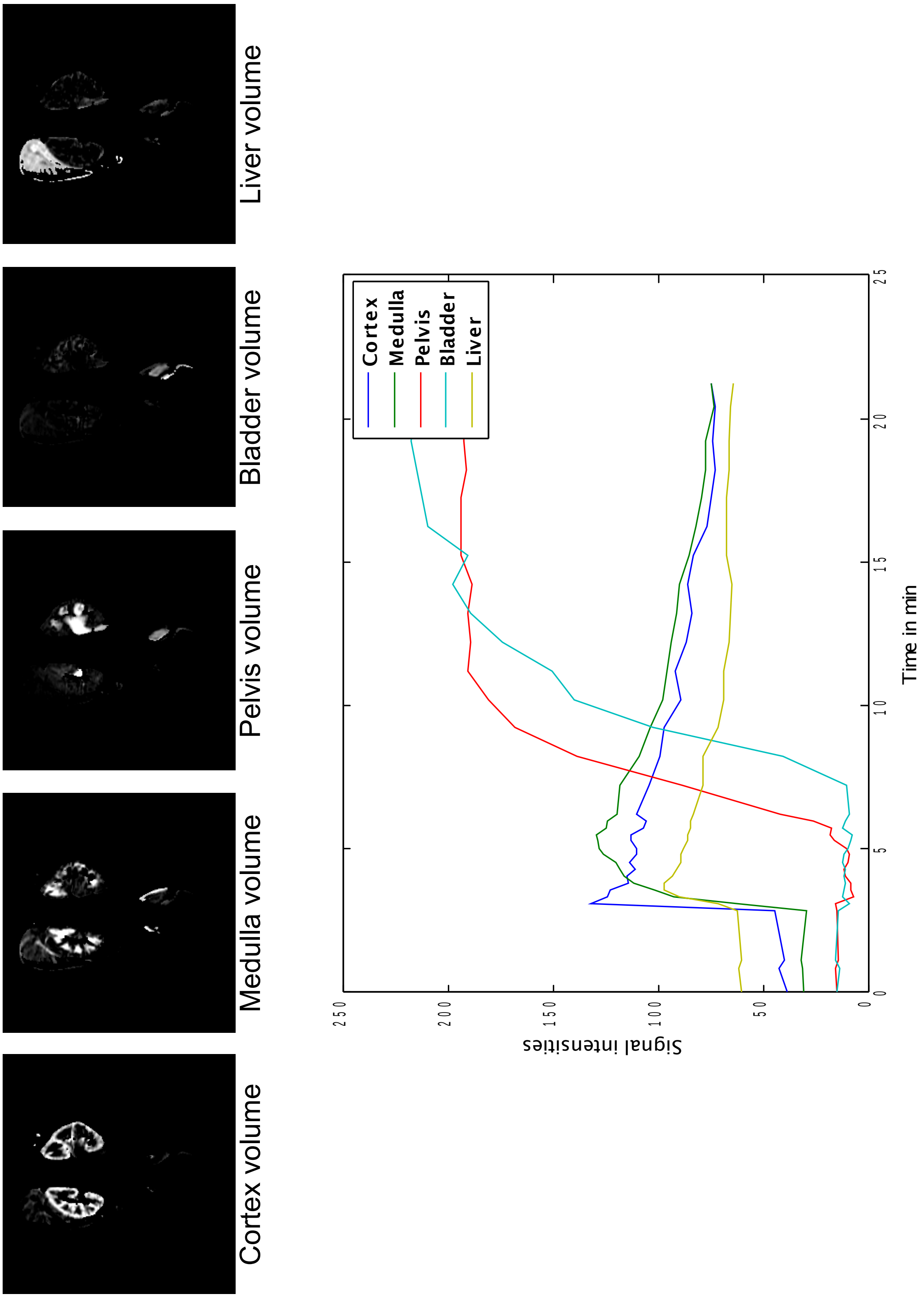


Figure 36: Factor volumes and Factor curves for data set 060721cpa1

The different phases of the renal enhancement pattern, reflecting the physiology of the kidneys, are clearly visible on the factor volumes and factor curves returned by FAMIS: pre-enhancement phase, vascular phase, tubular phase and excretory phase.

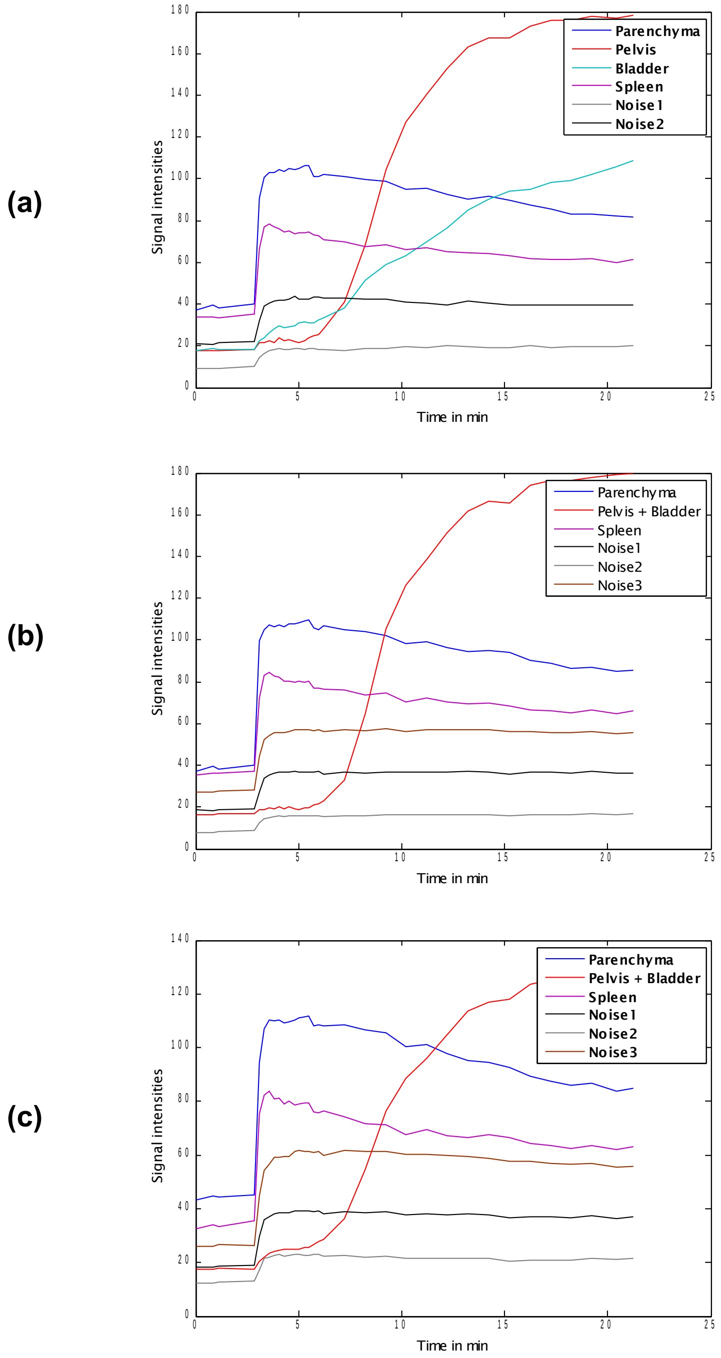


Figure 37: Representative TIC for data set 060721cpa1

- *From top to down: representative TIC output respectively from* **(a)** *a Kohonen Neural Network,* **(b)** *a Fuzzy K-means and* **(c)** *a EM clustering method.*
- *No distinct cortex and medulla components are visible as in FAMIS (Figure 36) but rather a mixture of both cortex and medulla in a parenchyma component.*
- *Clustering methods tend to regroup the quadrixels based on their signal amplitude.*

5.5 Discussion

For the majority of the data sets tested, the segmentation results provided by our proposed methods lie within one or two voxels of the ground truth, which means that they are close to the segmentation of our domain expert. However, there are few cases where our proposed clustering methods fail to give a result close to the ground truth ($H(A,B)>2$) due to the presence of the spleen or the liver in our segmentation results.

On average, our segmentation results seem to be closer to the ground truth for the right kidney than for the left kidney. This may be explained by two reasons:

- The spleen is a tissue that can raise problem to separate from the left kidney due to its similar time course with the renal cortex.
- In our data sets, the left kidney is the one which is the most frequently impaired and thus, also the one whose segmentation is more likely to be subject to variability.

KNN seems to be the method that performs worse. This is confirmed by Student *t*-tests that demonstrate a higher undirected average distance to the ground truth compared to FKM and FAMIS. It can be due to the inability for KNN to consider the uncertainty for a voxel to belong to different tissue types. On contrary, FKM, EM and FAMIS allow an operator to set a tissue membership threshold based on visual inspection, which is a more suitable approach to handle voxel uncertainties.

It must be noted that the quantitative performance evaluation used to assess our segmentation results must be considered carefully. Indeed, such an evaluation method judges the quality of our results compared to a ground truth, as if the ground truth was the “ideal” segmented image. Yet, it is expected that manual segmentations are subject to inter-observer and intra-observer variability, especially dealing with poorly functioning kidneys and partial volume effects. Another problem is that the ground truth is obtained by performing a manual segmentation mainly on a single volume at a certain time point whereas our segmentation approaches consider the kinetics of the whole dynamic volumetric sequence. Hence, our segmentation results are all the more different from the ground truth since the data sets are not perfectly corrected against motion (translation, rotation and deformation).

Qualitatively, FAMIS is the method that yields the most remarkable results: the factor volumes and factor curves returned by FAMIS may allow to assess the ability of the renal cortex and medulla to concentrate and excrete the contrast agent and thus, give insights into the physiology of the patient kidneys. Hence, if the desire for a clinician is to obtain within few minutes an exhaustive summary which can be physiologically interpretable from a huge 4-D data set, then FAMIS is the most promising approach. The main drawback is that it returns the spleen and the renal cortex in the same factor volume. This can be explained physiologically by the high vascularity of both tissues which dominantly enhance during the vascular phase of contrast agent uptake.

Finding a way for FAMIS to remove the spleen to obtain a factor volume that would contain only the renal cortex would be of great clinical interest because it would give the possibility to measure the functioning cortical volume. In the proposed approach, there is an attempt to exclude the spleen from analysis by using a high threshold Θ_I on the mean TIC. But this idea is not satisfying because it requires a careful selection of a threshold value by the operator and because it tends also to chop off part of the kidney tissue, especially the medulla. Another possible approach in the future is to combine both a clustering analysis and a factor analysis:

- A clustering analysis by FKM or EM first removes the spleen, which is proved to work well in this task because it is more sensitive to the difference in TIC amplitude.
- A factor analysis is subsequently applied on the data set completely devoid of splenic tissue and hence is able to return a pure renal cortical curve associated to a renal cortical volume that can be physiologically be interpretable by the clinician.

6 Conclusion

6.1 Results summary

A novel approach is presented for kidney segmentation in 4-D DCE-MRU based on the observation that the reproducible time-enhancement pattern of abdominal tissues is a strong feature to distinguish them.

Clustering techniques such as KNN, FKM and EM are applied to regroup the quadrixels based on their similarities and a FAMIS with novel constraints is applied to yield physiological factor volumes and associated factor curves.

Clustering techniques can return a renal parenchyma volume, but without distinguishing the renal cortex and renal medulla tissues, and a splenic tissue for most of the data sets. This may be due to the sensitivity of the approach to signal amplitudes.

One of the major benefits of FAMIS is that it can return the renal cortex and renal medulla in separate factor volumes. Consequently, it allows both separate and combined clinical analysis of cortical and medullary TIC, which may be of future importance in clinical renal diagnosis. But without the user interaction to remove the splenic tissue, FAMIS would return in a common factor volume both the renal cortex and the spleen. This may be due to the sensitivity of the approach to renal physiology: this common factor volume represents the vascular volume that comprises of the two abdominal tissues that are highly vascular and thus dominantly enhance during the arterial phase.

Our 4 proposed methods are quantitatively evaluated on 20 MR real patient data sets and achieve reasonably accurate results compared to the ground truth: the undirected average distance lies within 2 voxels on average.

Segmentation failure cases concern only clustering methods and are due to the presence of surrounding organs like spleen and liver in the results. On contrary, our FAMIS approach yields good segmentation results for all our data sets. Moreover, it provides factor curves with a clear visualization of the three physiological phases of the contrast agent uptake and thus, allow us to consider it for clinical routine.

6.2 Future work

Further work to demonstrate the clinical usefulness of FAMIS should be made: possible reliable parameters (like peak medullary signal intensity, cortico-medullary cross-over point, time between cortex peak and medulla peak, renal transit time...) could be derived from the factor volumes and associated factor curves to evaluate the functioning renal volume or assess the degree of renal obstruction. If the results were consistent with the expert examination, FAMIS would provide a powerful tool for renal diagnosis.

A method combining a cluster analysis to remove the splenic tissue before submitting the data set to FAMIS is a possible approach for a future work.

Moreover, our MR data sets are limited by temporal resolution (44 data points in TIC at most). Nowadays, new imaging techniques make possible to obtain more frequent sampling points in the TIC (more than 200). By analysing data sets where the time between each time point is shortened, it may be possible for our clustering methods to differentiate intrarenal compartments, and for our FAMIS to differentiate the spleen from the renal cortex.

Finally, our segmentation methods are not restricted to MRU only. It would be interesting to also investigate their performance in renal nuclear scintigraphy and DCE-MRI of tumours in various tissues.

Bibliography

[**1**] Dempster A, Laird N, Rubin D. *Maximum likelihood from incomplete data via the EM algorithm*. J. of the Royal Statistical Society (1977) 39 1: 1–38

[**2**] Barber DC. *The use of principal components in the quantitative analysis of gamma camera dynamic studies*. Phys. Med. Biol. (1980) 25 : 283-292

[**3**] Bezdek JC. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Kluwer Academic Publishers (1981) ISBN: 0306406713.

[**4**] Glazer GM, Axel L, Goldberg HI, Moss AA. *Dynamic CT of the normal spleen*. AJR Am. J. Roentgenol. (1981) 137 2: 343-346

[**5**] Di Paola R, Bazin JP, Aubry F, Aurengo A, Cavailloles F, Herry JY, Kahn E. *Handling of dynamic sequences in nuclear medicine*. IEEE Trans. Nucl. Sci. (1982) 29 : 1310-1321

[**6**] Cavailloles F, Bazin JP, Di Paola R. *Factor analysis in gated cardiac studies*. J. Nucl. Med. (1984) 25 : 1067-1079

[**7**] Houston AS. *The effect of apex-finding errors on factor images obtained from factor analysis and oblique transformation*. Phys. Med. Biol. (1984) 29 9: 1109–1116

[**8**] Nijran KS, Barber DC. *Towards automatic analysis of dynamic radionuclide studies using principal-components factor analysis*. Phys. Med. Biol. (1985) 30 : 1315-1325

[**9**] Nijran KS, Barber DC. *Factor analysis of dynamic function studies using a priori physiological information*. Phys. Med. Biol. (1986) 31 : 1107-1117

[**10**] Nakamura M, Suzuki Y, Kobayashi S. *A Method for Recovering Physiological Components from Dynamic Radionuclide Images Using the Maximum Entropy Principle: A Numerical Investigation*. IEEE Trans. Biom. Eng. (1989) 36 : 906-917

[**11**] Mirowitz SA, Brown JJ, Lee JK, Heiken JP. *Dynamic gadolinium-enhanced MR imaging of the spleen: normal enhancement patterns and evaluation of splenic lesions*. Radiology (1991) 179 3: 681-686

[**12**] Mirowitz SA, Gutierrez E, Lee JK, Brown JJ, Heiken JP. *Normal abdominal enhancement patterns with dynamic gadolinium-enhanced MR imaging*. Radiology (1991) 180 3: 637-640

[**13**] Bonnerot V, Charpentier A, Frouin F, Kalifa C, Vanel D, Di Paola R. *Factor analysis of dynamic MR imaging in predicting the response of osteosarcoma to chemotherapy*. Invest. Radiol. (1992) 27 : 847-855

[**14**] Semelka RC, Shoenut JP, Lawrence PH, Greenberg HM, Madden TP, Kroeker MA. *Spleen: dynamic enhancement patterns on gradient-echo MR images enhanced with gadopentetate dimeglumine*. Radiology (1992) 185 2: 479-482.

[**15**] Benali H, Buvat I, Frouin F, Bazin JP, Di Paola R. *A statistical model for the determination of the optimal metric in Factor Analysis of Medical Image Sequences (FAMIS)*. Phys. Med. Biol. (1993) 38 : 1065-1080

[**16**] Buvat I, Benali H, Frouin F, Bazin JP, Di Paola R. *Target apex-seeking in factor analysis of medical image sequences*. Phys. Med. Biol. (1993) 38 : 123-138

[**17**] Frouin F, Cinotti L, Benali H, Buvat I, Bazin JP, Millet P, Di Paola R. *Extraction of functional volumes from medical dynamic volumetric data sets*. Comput. Med. Imaging Graph. (1993) 17 4-5: 397-404

[**18**] Hendrick RE, Haacke EM. *Basic physics of MR contrast agents and maximization of image contrast*. J. Magn. Reson. Imaging (1993) 3 : 137

[**19**] Benali H, Buvat I, Frouin F, Bazin JP, Di Paola R. *Foundations of factor analysis of medical image sequences*. Image and Vision Computing (1994) 12 : 375-385

[**20**] Rabushka LS, Kawashima A, Fishman EK. *Imaging of the spleen: CT with supplemental MR examination*. Radiographics (1994) 14 2: 307-332

[**21**] Zagdansky AM, Sigal R, Bosq J, Bazin JP, Vanel D, Di Paola R. *Factor analysis of medical image sequences in MR of head and neck tumors*. Am. J. Neuroradiol. (1994) 15 : 1359-1368

[**22**] Wu HM, Hoh CK, Choi Y, Schelbert HR, Hawkins RA, Phelps ME, Huang SC. *Factor analysis for extraction of blood time-activity curves in dynamic FDG-PET studies*. J. Nucl. Med. (1995) 36 : 1714-1722

[**23**] Taylor J, Summers PE, Keevil SF, Saks AM, Diskin J, Hilton PJ, Ayers AB. *Magnetic resonance renography: optimisation of pulse sequence parameters and Gd-DTPA dose, and comparison with radionuclide renography*. Magnetic Resonance Imaging (1997) 15 6: 637-649

[**24**] Knesplova L, Krestin GP. *Magnetic resonance in the assessment of renal function*. Eur. Radiol. (1998) 8 : 201-211

[**25**] Nolte-Ernsting CC, Bucker A, Adam GB, Neuerburg JM, Jung P, Hunter DW, Jakse G, Gunther RW. *Gadolinium-enhanced excretory MR urography after low-dose diuretic injection: comparison with conventional excretory urography*. Radiology (1998) 209 : 147-157

[**26**] Samal M, Nimmon CC, Britton KE, Bergmann H. *Relative renal uptake and transit time measurements using functional factor images and fuzzy regions of interest*. Eur. J. Nucl. Med. (1998) 25 : 48-54

[**27**] Tomaru Y, Inoue T, Oriuchi N, Takahashi K, Endo K. *Semi-automated renal region of interest selection method using the double-threshold technique: inter-operator variability in quantitating 99mTc-MAG3 renal uptake*. Eur. J. Nucl. Med. (1998) 25 1: 55-59

[**28**] Bruyant PP, Sau J, Mallet JJ. *Noise removal using factor analysis of dynamic structures: application to cardiac gated studies*. J. Nucl. Med. (1999) 40 : 1676-1682

[**29**] Buvat I, Hapdey S, Benali H, Todd-Pokropek A, Di Paola R. *Spectral factor analysis for multi-isotope imaging in nuclear medicine*. Information Processing in Medical Imaging (1999) : 442-447

[**30**] Chuang KH, Chiu MJ, Lin CC; Chen JH. *Model-Free Functional MRI Analysis Using Kohonen Clustering Neural Network and Fuzzy C-Means*. IEEE Trans. Med. Imaging (1999) 18 12: 1117 - 1128

[**31**] Haacke EM, Brown RW, Thompson MR, Venkatesan R. *Magnetic resonance imaging : physical principles and sequence design*. Wiley-Liss (1999) ISBN: 0471351288.

[**32**] Haykin S. *Neural Networks: a comprehensive foundation*. Prentice Hall (1999) ISBN: 0-13-273350-1.

[**33**] Sitek A, Di Bella EV, Gullberg GT. *Factor Analysis of Dynamic Structures in Dynamic SPECT Imaging Using Maximum Entropy*. IEEE Trans. Nucl. Sci. (1999) 46 6: 2227-2232

[**34**] Stark DD, Bradley WG. *Magnetic Resonance Imaging Vol1 3rd Edition*. Mosby Inc, St. Louis (1999) ISBN: 0815185189.

[**35**] Duda RO, Stork DG, Hart PE. *Pattern Classification*. Wiley-Interscience (2000) ISBN: 978-0471056690.

[**36**] Sitek A, Di Bella EV, Gullberg GT. *Factor analysis with a priori knowledge-application in dynamic cardiac SPECT*. Phys. Med. Biol. (2000) 45 9: 2619-2638

[**37**] Stanisz GJ, Henkelman RM. *Gd-DTPA relaxivity depends on macromolecular content*. Magn. Reson. Med. (2000) 44 : 665-667

[**38**] Boykov Y, Lee VS, Rusinek H, Bansal R. *Segmentation of Dynamic N-D Data Sets via Graph Cuts Using Markov Models*. MICCAI (2001) Lecture Notes in Computer Science 2208 : 1058-1066

[**39**] de Priester JA, Kessels AG, Giele EL, den Boer JA, Christiaans MH, Hasman A, van Engelshoven JM. *MR renography by semiautomated image analysis: performance in renal transplant recipients*. J. Magn. Reson. Imaging (2001) 14 2: 134-140

[**40**] Krier JD, Ritman EL, Bajzer Z, Romero JC, Lerman A, Lerman LO. *Noninvasive measurement of concurrent single-kidney perfusion, glomerular filtration, and tubular function*. Am. J. Physiol. Renal Physiol. (2001) 281 4: F630-F638

[**41**] Martel AL, Allder SJ, Delay GS, Morgan PS, Moody AA. *Perfusion MRI of infarcted and noninfarcted brain tissue in stroke: a comparison of conventional hemodynamic imaging and factor analysis of dynamic studies*. Invest. Radiol. (2001) 36 7: 378-385

[**42**] Martel AL, Moody AR, Allder SJ, Delay GS, Morgan PS. *Extracting parametric images from dynamic contrast-enhanced MRI studies of the brain using factor analysis*. Med. Image Anal. (2001) 5 1: 29-39

[**43**] Coulam CH, Bouley DM, Sommer FG. *Measurement of renal volumes with contrast-enhanced MRI*. J. Magn. Reson. Imaging (2002) 15 2: 174-179

[**44**] Janier MF, Mazzadi AN, Lionnet M, Frouin F, Andre-Fouet X, Cinotti L, Revel D, Croisille P. *Factor analysis of medical image sequences improves evaluation of first-pass MR imaging acquisitions for myocardial perfusion*. Acad. Radiol. (2002) 9 1: 26-39

[**45**] Sitek A, Di Bella EV, Gullberg GT, Huesman RH. *Removal of liver activity contamination in teboroxime dynamic cardiac SPECT imaging with the use of factor analysis*. J. Nucl. Cardiol. (2002) 9 2: 197-205

[**46**] Sitek A, Gullberg GT, Huesman RH. *Correction for ambiguous solutions in factor analysis using a penalized least squares objective*. IEEE Trans. Med. Imaging. (2002) : 216-225

[**47**] Sun Y, Moura JMF, Yang D, Ye Q, Ho C. *Kidney segmentation in MRI sequences using temporal dynamics*. IEEE International Symposium on Biomedical Imaging (2002) : 98 - 101

[**48**] Tsagaan B, Shimizu A, Kobatake H, Miyakawa K. *An Automated Segmentation Method of Kidney Using Statistical Information*. MICCAI (2002) Lecture Notes in Computer Science 2488 : 556–563

[**49**] Martel AL, Fraser D, Delay GS, Morgan PS, Moody AA. *Separating arterial and venous components from 3D dynamic contrast-enhanced MRI studies using factor analysis*. Magn. Reson. Med. (2003) 49 5: 928-933

[**50**] Rohrschneider WK, Haufe S, Clorius JH, Tröger J. *MR to assess renal function in children*. Eur. Radiol. (2003) 13 : 1033-1045

[**51**] Teh HS, Ang ES, Wong WC, Tan SB, Tan AG, Chng SM, Lin MB, Goh JS. *MR Renography Using a Dynamic Gradient-Echo Sequence and Low-Dose Gadopentetate Dimeglumine as an Alternative to Radionuclide Renography*. AJR Am. J. Roentgenol. (2003) 181 2: 441-450

[**52**] Bruno S, Remuzzi G, Ruggenenti P. *Transplant renal artery stenosis*. J. Am. Soc. Nephrol. (2004) 15 : 134-141

[**53**] de Bazelaire CM, Duhamel GD, Rofsky NM, Alsop DC. *MR Imaging Relaxation Times of Abdominal and Pelvic Tissues Measured in Vivo at 3.0 T: Preliminary Results*. Radiology (2004) 230 3: 652-659

[**54**] Ell P, Gambhir S. *Nuclear medicine in clinical diagnosis and treatment*. Churchill Livingstone (2004) ISBN: 0443073120.

[**55**] Frouin F, Delouche A, Raffoul H, Diebold H, Abergel E, Diebold B. *Factor analysis of the left ventricle by echocardiography (FALVE): a new tool for detecting regional wall motion abnormalities*. Eur. J. Echocardiography (2004) 5 : 335-346

[**56**] Kahn E, Lizard G, Dumas D, Frouin F, Menetrier F, Stoltz JF, Todd-Pokropek A. *Analysis of Fluorescent MRI Contrast Agent Behavior in the Liver and Thoracic Aorta of Mice*. Anal. Quant. Cytol. Histol. (2004) 26 4: 233-238

[**57**] Subramaniam M, Mizzi A, Roditi G. *Magnetic resonance angiography in potential live renal donors: a joint radiological and surgical evaluation*. Clin. Radiol. (2004) 59 : 335-341

[**58**] El Fakhri G, Sitek A, Guérin S, Kijewski MF, Di Carli MF, Moore SC. *Quantitative Dynamic Cardiac (82)Rb PET Using Generalized Factor and Compartment Analyses*. J. Nucl. Med. (2005) 46 8: 1264-1271

[**59**] Jones RA, Easley K, Little SB, Scherz H, Kirsch AJ, Grattan-Smith JD. *Dynamic contrast-enhanced MR urography in the evaluation of pediatric hydronephrosis: Part 1, functional assessment*. AJR Am. J. Roentgenol. (2005) 185 6: 1598-1607

[**60**] Rao M, Stough J, Chi YY, Muller K, Tracton G, Pizer SM, Chaney EL. *Comparison of human and automatic segmentations of kidneys from CT images*. Int. J. Radiat. Oncol. Biol. Phys. (2005) 61 3: 954-960

[**61**] van den Dool SW, Wasser MN, de Fijter JW, Hoekstra J, van der Geest RJ. *Functional Renal Volume: Quantitative Analysis at Gadolinium-enhanced MR Angiography-Feasibility Study in Healthy Potential Kidney Donors*. Radiology (2005) 236 : 189-195

[**62**] Grattan-Smith JD, Jones RA. *MR urography in children*. Pediatric Radiology (2006) 36 : 1119-1132

[**63**] Grenier N, Hauger O, Cimpean A, Pérot V. *Update of Renal Imaging*. Semin. Nucl. Med. (2006) 36 : 3-15

[**64**] Kirsch AJ, Grattan-Smith JD, Molitierno Jr JA. *The role of magnetic resonance imaging in pediatric urology*. Curr. Opin. Urol. (2006) 16 : 283-290

[**65**] Koh HK, Shen W, Shuter B, Kassim AA. *Segmentation of Kidney Cortex in MRI studies using a Constrained Morphological 3D H-maxima Transform*. Ninth International Conference on Control, Automation, Robotics & Vision (2006)

[**66**] Michoux N, Vallée J-P, Pechère-Bertschi A, Montet X, Buehler L, Van Beers BE. *Analysis of contrast-enhanced MR images to assess renal function*. Magn. Reson. Mater. Phy. (2006) 19 : 167-179

[**67**] Wismuller A, Meyer-Baese A, Lange O, Reiser MF, Leinsinger G. *Cluster Analysis of Dynamic Cerebral Contrast-Enhanced Perfusion MRI Time-Series*. IEEE Trans. Med. Imag. (2006) 25 1: 62-73

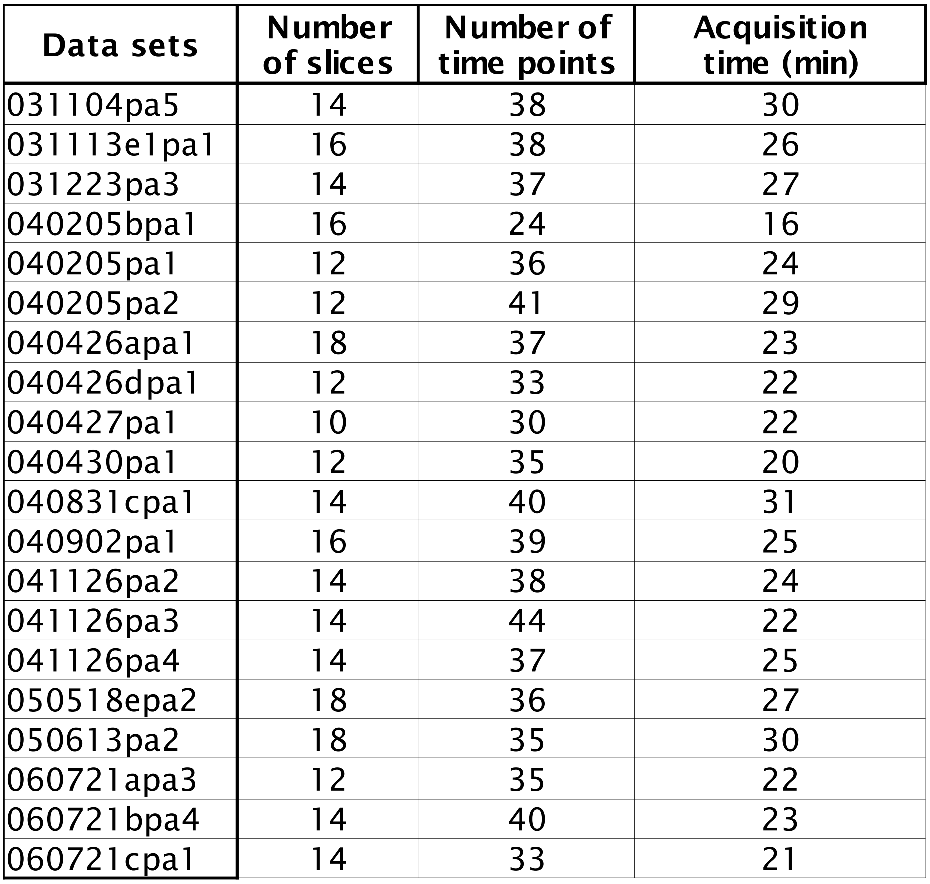
[**68**] Lin S-P, Brown JJ, MD. *MR Contrast Agents: Physical and Pharmacologic Basics*. J. Magn. Reson. Imaging (2007) 25 : 884-899

[**69**] McRobbie DW, Moore EA, Graves MJ, Prince MR. *MRI from Picture to Proton*. Cambridge University Press (2007) ISBN: 9780521683845.

A Appendix

The denomination of our 20 patient MR data sets, the number of slices contained in their 3-D volumes, their number of time points (that is, the number of 3-D volumes) and their acquisition time are given in Figure 38.

Figure 39 and Figure 40 are illustrations of a typical MR data set.



Data sets	Number of slices	Number of time points	Acquisition time (min)
031104pa5	14	38	30
031113e1pa1	16	38	26
031223pa3	14	37	27
040205bpa1	16	24	16
040205pa1	12	36	24
040205pa2	12	41	29
040426apa1	18	37	23
040426dpa1	12	33	22
040427pa1	10	30	22
040430pa1	12	35	20
040831cpa1	14	40	31
040902pa1	16	39	25
041126pa2	14	38	24
041126pa3	14	44	22
041126pa4	14	37	25
050518epa2	18	36	27
050613pa2	18	35	30
060721apa3	12	35	22
060721bpa4	14	40	23
060721cpa1	14	33	21

Figure 38: Characteristics of our 20 MR data sets

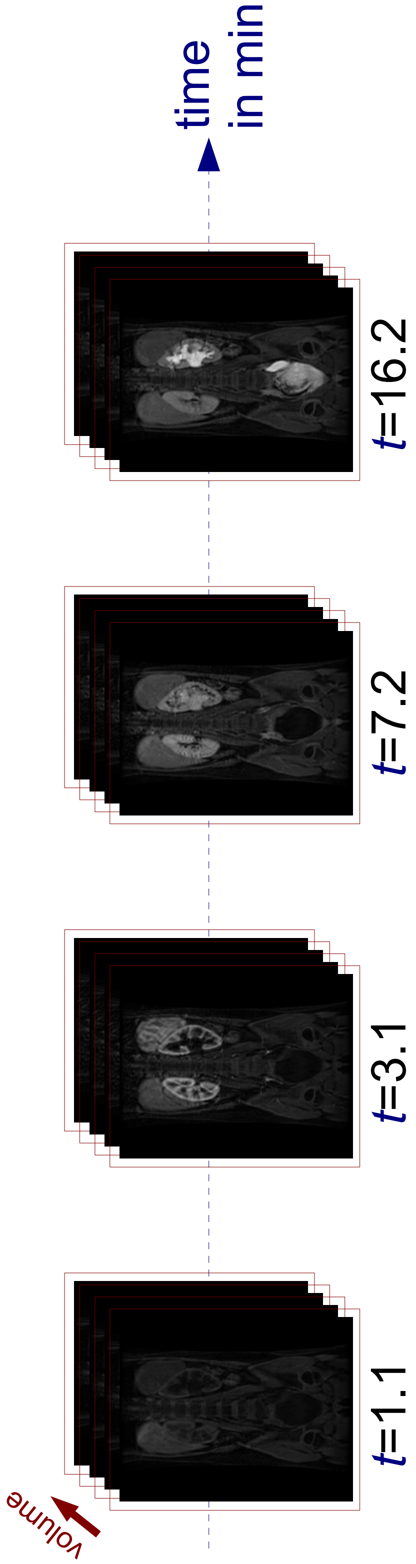


Figure 39: Illustration of a sequence of 3-D images (4-D data set)

This is an illustration of data set 060721cpa1 which is a sequence of 33 volumetric images made of 14 image slices during an acquisition time of 21 min.

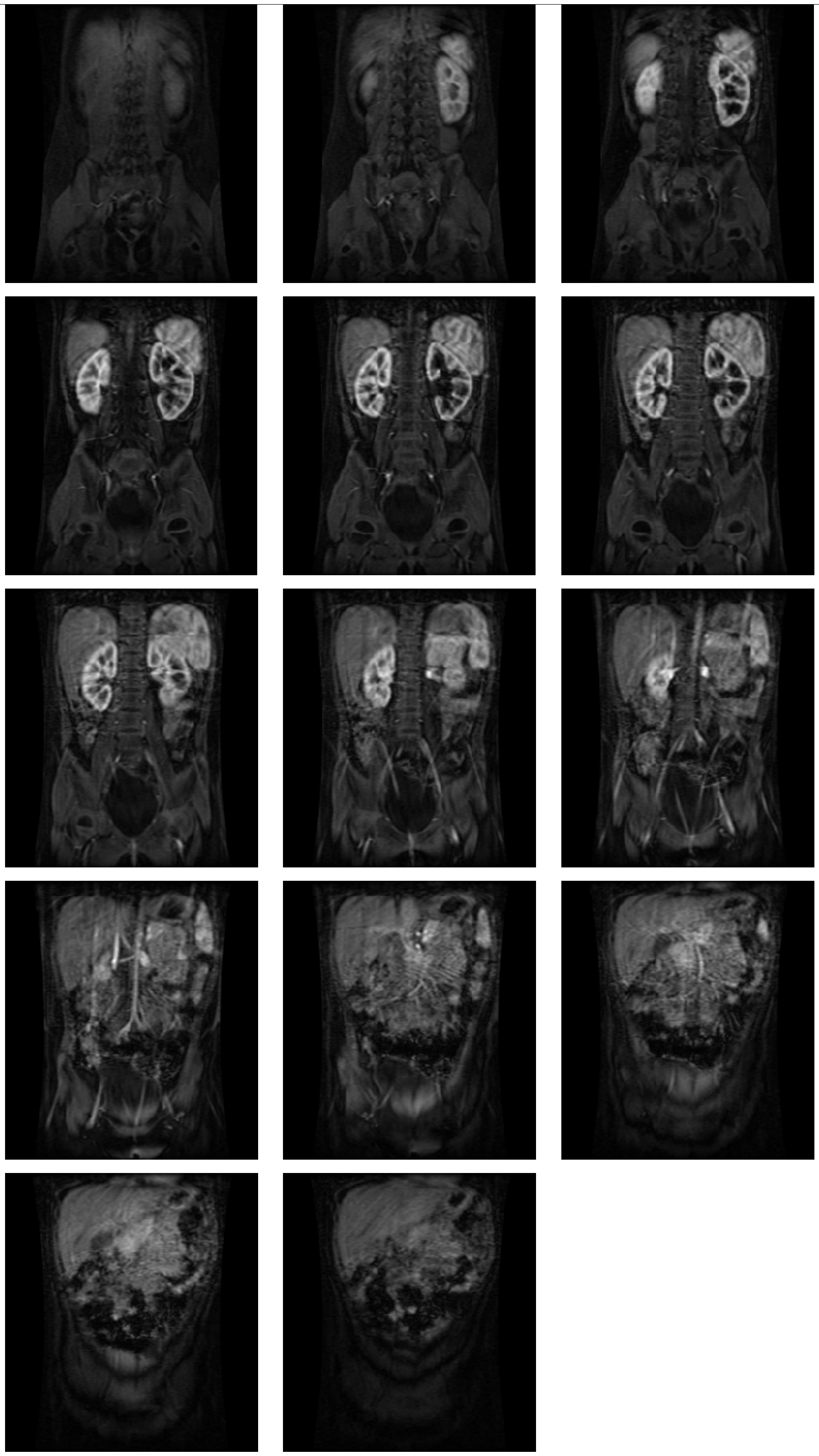


Figure 40: Illustration of 1 volumetric image made of 14 slices

- *From top to down and left to right: 14 slices, from most posterior to most anterior, of 1 volumetric image chosen when the cortico-medullary differentiation is the best.*
- *Data set 060721cpa1 at t=3.1 min.*