UNIVERSITAT POLITÈCNICA DE VALÈNCIA

ESCUELA TÉCNICA SUPERIOR INGENIEROS INDUSTRIALES VALENCIA

TRABAJO FIN DE GRADO EN INGENIERÍA BIOMÉDICA

DEVELOPMENT OF A PREDICTION MODEL FOR TUMOR DETECTION IN PROSTATE USING MULTIVARIATE IMAGE ANALYSIS FROM PERFUSION AND DIFFUSION MR SEQUENCES

ZIQING WANG

CO-SUPERVISERIC AGUADO SARRIÓ, ALBERTO JOSÉ FERRER RIQUELM

SUPERVISOJOSÉ MANUEL PRATS MONTALBÁN

Academic year: 2015-16

Acknowledgment

I can still recall my first day in class as if it were yesterday. Yet when open my eyes, here I am, writing acknowledgment as finishing touch on my thesis. Writing this thesis, as well as studying in Spain has had a huge impact on me. I would like to reflect on all the people who have supported me throughout my time in this great university. Turns out that there are too many people to give thanks to, but too little space.

First I would like to thank my family, my parents and my grandparents. I know from my heart that you would always be there for me. Without your support, both financially and spiritually, I could never go this far.

I am also grateful to all professors whom I have encountered in four years. Among them, I would like to thank especially Alberto José Ferrer, Rafael Ors, María Purificación Lisón, Isabel María Morera, and so on, for your patience, encouragement, and valuable guidance.

Other than that, I am extremely grateful and indebted to my tutors José Manuel Prats and Eric Aguado. I could not thank you both enough for your shared expertise, help and guidance. You definitely provided me with all the tools and materials that I needed to complete this thesis.

Last but not least, I want to thank all my friends, those back at home and those I made here. Thank you for your sympathetic ear and wonderful companionship. I would particularly like to single out one of my dearest friends, Alejandro Hueso, for everything. With you, I believe that no words are necessary after all the things and course work we have been through together.

I would cherish all these precious memories. Thank you all so much!

Ziqing Wang

Valencia, July 3, 2016.

Abstract

Nowadays more and more physiological models are developed and applied to extract imaging biomarkers from medical images, which have been used for cancer and lesion detection in different organs, prostate, breast, lung, brain, etc. So the current trending is to study the potential of the combination of these emerged biomarkers in diagnosis, increasing the need for image coregistration.

This thesis mainly consists of aligning MR (Magnetic Resonance) prostate perfusion and diffusion sequences using software Horos. After the image coregistration, same biomarkers established by previous studies (José Manuel Prats and others, 2013, Eric Aguado and others, 2014) were extracted using Multivariate Curve Resolution (MCR), and later analyzed with Partial Least Square regression (PLS) model. The obtained results show a good prediction with f-score 0.8144. Similar outcome has also been achieved with leave one out cross validation (LOOCV). As a preliminary work, further studies are recommended and necessary to evaluate better the precision and to extend the work with more depth.

Key words: MIA, MCR, PLS, MR, Perfusion, Diffusion, Image coregistration

Resumen

Hoy en día hay cada vez más modelos fisiológicos desarrollados y aplicados para extraer biomarcadores de imágenes médicas. Estos biomarcadores son usados mayormente en detectar lesión y cáncer en diferentes órganos, próstata, mama, pulmón, cerebro, etc. Por lo cual la tendencia actual se convierte en estudiar el potencial de la combinación de estos biomarcadores en el diagnóstico, aumentando la necesidad de corregistro de imagen.

Esta tesis consiste principalmente en la alineación de RM (resonancia magnética) secuencias de perfusión y difusión utilizando el software de Horus. Después de la alineación, los mismos biomarcadores establecidos por investigaciones previas (José Manuel Prats y otros, 2013, Eric Aguado y otros, 2014) son extraídos usando Curva de Resolución Multivariante (MCR), y luego analizados con regresión de mínimos cuadrados parciales (PLS). Los resultados obtenidos muestran una buena predicción con f-score 0.8144. Con *leave one out* validación cruzada (LOOCV) también se obtiene resultado parecido. Como un trabajo preliminar, estudios adicionales y más profundos son recomendados y necesarios para evaluar mejor la precisión y extenderlo con más profundidad.

Palabras Clave: MIA, MCR, PLS, MR, Perfusión, Difusión, Corregistro de imagen

Resum

Hui en dia hi ha cada vegada més models fisiològics desenrotllats i aplicats per a extraure biomarcadors d'imatges mèdiques. Estos biomarcadors són usats majorment per a detectar lesió i càncer en diferents òrgans, pròstata, mamella, pulmó, cervell, etc. Per la qual cosa la tendència actual es convertix a estudiar el potencial de la combinació d'estos biomarcadors en el diagnòstic, augmentant la necessitat de corregistre d'imatge.

Esta tesi consistix principalment en l'alineació de RM (ressonància magnètica) seqüències de perfusió i difusió utilitzant el programa Horos. Després de l'alineació, els mateixos biomarcadores establits per investigacions prèvies (José Manuel Prats i altres, 2013, Eric Aguado i altres, 2014) són extrets usant Corba de Resolució Multivariant (MCR) , i després analitzats amb regressió de mínims quadrats parcials (PLS). Els resultats obtinguts mostren una bona predicció amb f-score 0.8144. Amb leave one out validació encreuada (LOOCV) també s'obté resultat paregut. Com un treball preliminar, estudis addicionals i més profunds són recomanats i necessaris per a avaluar millor la precisió i estendre-ho amb més profunditat.

Paraules clau: MIA, MCR, PLS, MR, Perfusió, Difusió, Corregistro d'imatge

Table of contents

DISSERTATION DOCUMENT

Chapter 1. INTRODUCTION

1.1. Objectives

The main object of this thesis is to align the perfusion and diffusion Magnetic Resonance sequences using the software Horos, to investigate whether combined biomarkers of these two sequences provide a good prediction for prostate tumor.

Once the alignment is finished, the aligned sequences are analyzed through multivariate analysis, specifically Multivariate Curve Resolution (MCR) and Partial Least Squares regression (PLS). As the result of this study, a predictive model for tumor detection in prostate is developed and capable of explaining associated physiologic phenomena in those sequences, and meanwhile extracting clinical information for diagnostics.

1.2. Prostate Cancer

Carcinoma of the prostate, or generally called as prostate cancer, is the most common neoplasia in Europe, exceeding lung and colorectal cancer (Jemal A and others, 2008). In Spain, the situation is no different, with more than 25.000 cases diagnosed annually, 21% of tumors among men.

Early prostate cancer usually does not cause clear symptoms, which makes early diagnose more difficult than others if the patient does not schedule regular physical examination. Sometimes when it does cause symptoms, the patient may undergo urinary dysfunction such as frequent and/or painful urination, increased urination at night, and blood in the urine. These symptoms are due to the fact that the prostate gland surrounds a part of the urethra, proximal to the bladder. And because the secretions from the prostate are also included in semen, prostate cancer may result in problems with erection or painful ejaculation (Miller DC and others, 2003).

Although most prostate cancers are slow growing, some can grow more quickly. When the cancer cells penetrate through prostate gland to adjacent tissue, it may spread to other parts of the body causing additional different symptoms. This often happens with advanced prostate cancer, probably in bones and lymph nodes. By that time the patient would feel pain in the pelvis, vertebrae, ribs or leg weakness should the cancer reach the spine and compress the spinal cord.

There are three primary risk factors: race, family history and age. The prostate cancer is very rare in men younger than 45, and the average age at the time of diagnosis is 70 (Hankey BF and others, 1999). Ethnically, the risk is relatively higher in black people, with a prevalence rate of 275 per 100,000. However, the risk is doubled for men who have one first-degree family member with prostate cancer compared to those without. And the risk is increased between 5 to 11 times when there are two or more first-degree family members affected. On the other hand, mutations in genes like SNPs, HPC1, BRCA1 and BRCA2 have been linked to prostate cancer.

An intriguing fact has been found considering the prostate carcinoma. Most of the prostate cancer originates in the peripheral area rather than the central zone, while the benign prostatic hyperplasia normally affects the transitional or periurethral zone of the prostate. Yet there is no evidence to correlate these two diseases.

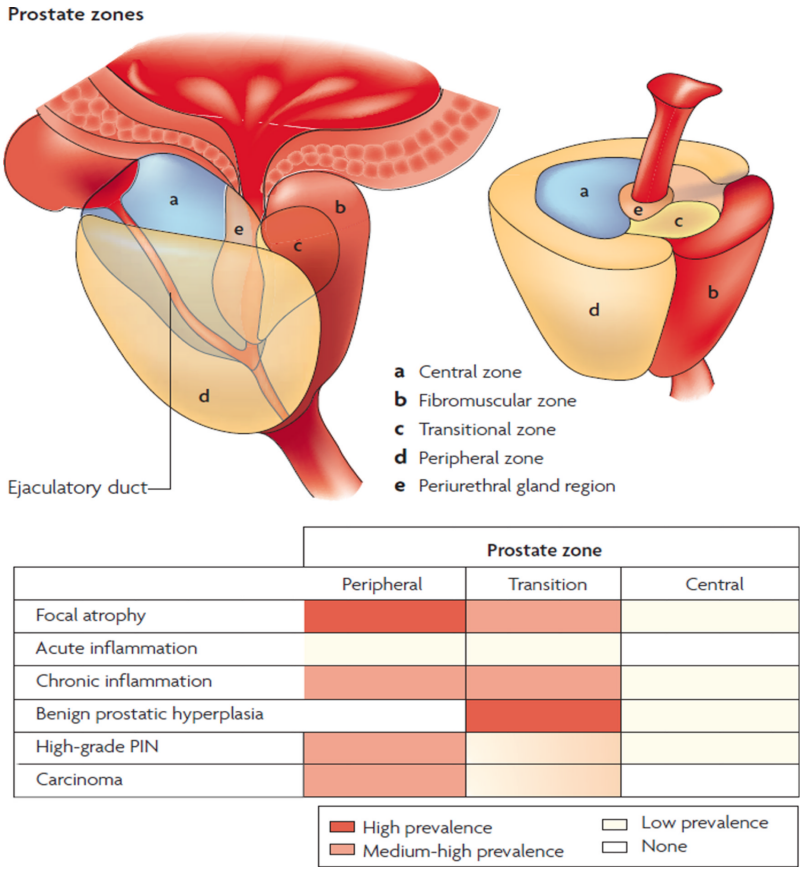


Figure 1. Different areas of prostate with its' prevalence of common diseases. (David Marti Aguado, 2014)

As for diagnosis, only through biopsy can the prostate cancer be fully confirmed. Initial diagnosis can be conducted with less invasive techniques such as digital rectal examination (DRE), prostate-specific antigen (PSA), and transrectal ultrasonography. The PSA is a marker of prostatic tissue usually found higher in cancer cells, but it only serves as a complementary result since it is also the consequence of prostatitis or other benign diseases. Not to mention that some of the prostate cancer cases even give normal results. Generally if the rectal examination turned out to be suspicious and a PSA superior to 4 ng/ml, it would be sufficient criteria for a prostatic biopsy. It is recommended to all men over 50 years to have a DRE and PSA every year. However, studies have shown that this recommendation for population screening of prostate cancer is not convenient owing to the risks of over-diagnosis and over-treatment (Andriole GL and others, 2009; Schröder FH and others, 2009).

1.3. Prostate Imaging

Thanks to the great advances in medical imaging, the diagnosis of prostate cancer is extended to non-invasive, painless method using also biomarkers. An imaging biomarker is defined as an extracted feature of medical images, which can be measured objectively and yields quantitative information. Nowadays Magnetic Resonance Imaging (MRI) has become one of the most frequently applied medical imaging techniques in diagnosing prostate cancer, because of its versatility, wide range of image contrasts, especially its capacity to combine the spatial/temporal resolutions. On the other hand, it is rather safe compared to other techniques based on ionizing radiation.

According to different target organs, corresponding tissue property is weighted into frequency of the magnetic field. Then depending on the choices of various contrast media and configurable parameters during image acquisition, along with the help of multi-parameter techniques or image post-processing, the following information can be revealed in MRI images:

1. Anatomical information obtained from T2-weighted MRI images.

   Prostate cancer in T2-weighted MRI images is typically represented as a focal lesion of low signal intensity. Although it has been widely used for oncologists as the first evidence to detect prostate cancer before treatment, this modality is far from providing satisfactory sensitivity and specificity (Quint LE and others, 1991). Under these circumstances, both perfusion and diffusion sequences have been investigated and applied to improve the diagnostic performance of MRI.

1.3.1. Perfusion

2. Information of angiogenesis, also known as perfusion, through dynamic sequences acquired in T1-weighted MRI images, after Dynamic Contrast-Enhancement (DCE).

   Angiogenesis is a biological process associated to tissues with increased nutrient and oxygen demand resulting in the formation of new vessels, and thus increase the blood perfusion. Therefore it is commonly present in tumor cells but rare in

healthy subjects. The quantification of the tumor angiogenesis measurements has shown great potential in applications not only to differentiate prostate cancer from noncancerous tissue and gradate tumor stages, but also to evaluate therapeutic response early after treatment onset (Jackson ASN and others, 2009; Leach MO and others, 2005).

In order to attain quantitative measurements, one of the most popular approaches is the mathematical pharmacokinetic model that quantifies the variations in the local concentration of the contrast medium over time. Specifically saying, it characterizes the intensity versus time curves per pixel, providing physiologically meaningful parameters. Of critical importance is the appropriate selection of the arterial input function (AIF) based on the tumor-feeding artery (Figure 2). If represented in graphs, it would be composed of a baseline, an abrupt positive peak, and a fast decay.

The next step is to apply mathematical analysis to the tissue enhancement curves on a pixel-by-pixel basis. At last, the quantitative biomarkers K_{ep}, K^{trans}, V_e and V_p are obtained by curve fitting algorithms, also for each pixel. The general equation for the pharmacokinetic model is (Tofts PS and others, 1999):

$$C(x,y,t) = K^{trans} \int_0^t C_{AIF}(\tau) e^{-kep^{\tau}} \, d\tau \qquad (1)$$

where **C(x,y,t)** is the tissue enhancement curve at (x,y) pixel, **C_{AIF}(t)** the AIF enhancement curve, and **K_{ep}** the washout constant.

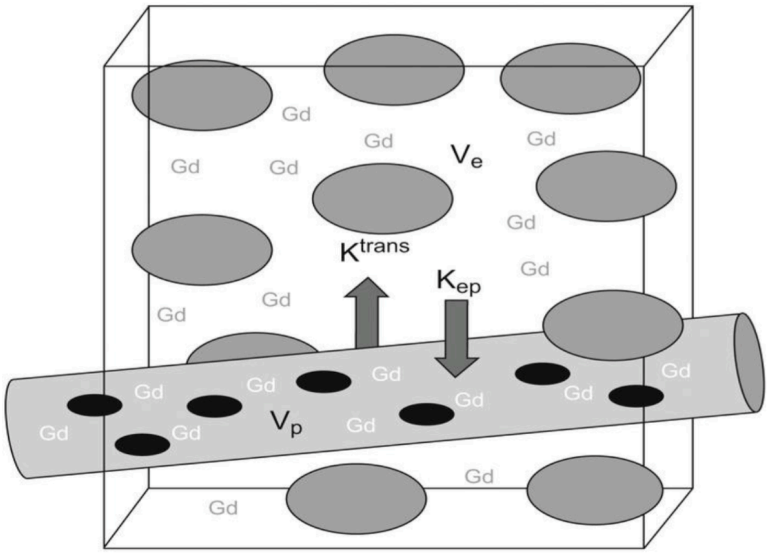


Figure 2. The four biomarkers extracted from the pharmacokinetic model. K_{ep}: coefficient of transference between plasma and extravascular extracellular space (EES); K^{trans}: transfer constant, V_e: volume fraction of the EES; V_p: volume fraction of plasma in the intravascular compartment. (David Marti Aguado, 2014)

These four biomarkers are of high correlation with the angiogenesis. And by representing them numerically or with regionally colored parametric maps, intratumoral heterogeneity of the vascular distribution can easily be analyzed (Barrett T and others, 2007).

Nevertheless, to determine this function requires visual inspection and manual selection of the region of interest (ROI). This process inevitably introduces user-dependent bias into the analysis, making almost impossible the reproducibility and standardization of different cases. Moreover, series of assumptions about the tissue vascular environment have to be made unless *a priori* knowledge is provided. All these limitations have prevented the widespread use of this model in clinical practice.

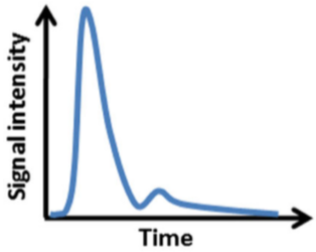


Figure 3. Dynamic patterns (AIF) of dynamic contrast-enhanced magnetic resonance images. (José Manuel Prats Montalbán and others, 2013)

An alternative is referring to Multivariate Image Analysis (MIA), particularly Multivariate Curve Resolution – Alternating Least Squares model (MCR-ALS). The idea behind MCR is that the pixel enhancement curve is now considered as a linear combination of various dynamic behaviors. Additionally, physiological interpretability is improved thanks to the lack of orthogonality restriction and the introduction of *a priori* knowledge to the model.

To compare with the mathematical pharmacokinetic model, the MCR-ALS is faster and easier to reproduce the results, no expert knowledge needed. Besides, not like PCA, it allows imposing constraints like non-negativity.

Therefore the MCR-ALS model is chosen and used to extract biomarkers in this thesis, and will be more thoroughly described in the next chapter.

1.3.2. Diffusion

3. Information about molecular movements from diffusion images.

Apart from angiogenesis, cellularization is another non-negligible sign for a tumor. They are the very first steps in oncogenesis, thus their presence usually confirms the existence of an early tumor. One possible way to look into these two biological processes is through studying the tissue local diffusion, resulted from the thermal agitation of the water molecules inside the body.

Diffusion is defined as a process in which free molecules move randomly in space. Inside of human body, cellular structures stand as barriers to diffusion of free water molecules. When the tissue is highly vascularized, free molecules can move at a high velocity within the vessels. Yet when the tissue is highly cellularized, free molecules are greatly restricted to movement since the interstitial space is drastically decreased.

The MRI is the only modality that permits the calculation and visualization of this process in vivo directly from the molecule movements (L. Martí-Bonmatí and others, 2010). This technique, Diffusion-Weighted Magnetic Resonance Imaging (DW-MRI), has the advantage of providing high-resolution images. Cellularized tissue (tumor) would appear with higher signal intensity, in accordance with its proliferation and aggressiveness.

The weighted magnetic resonance signal is very sensitive to water molecules movement. Due to the thermic agitation of water molecules, the relaxation of the spines (loss of synchronism) is accelerated. And the repetitive displacements of water make the nuclear spines spread to areas where the magnetic field would alter from the original, causing a modulation of the relaxation frequency.

The image acquisition of the DW-MRI depends on the configurations of the equipment and a key parameter, b-value (s/mm^2). It is essential because the image signal decreases as the acquired b-value increases, informing about the degree of

enhancement in diffusion. Furthermore, this attenuation changes with different characteristics of the tissue, being stronger regarding angiogenesis and weaker with cellularization. That is to say distinct kinds of tissue would show varied signal attenuation under the same b-value.

Normally the recommended rank for b-value is from 100 s/mm^2 to 800 s/mm^2, but for lesion detection and characterization purposes, 5 values (0, 50, 200, 400, 1000) are always applied in daily practice to study the attenuation of the signal.

In general terms, measured parameters of diffusion state the effective displacement of water molecule during a time interval (D. Le Bihan and others, 1988). It is assumed that at the initial instant molecules are concentrated together at one certain point. After a time interval, without outside forces, they would expand in three dimensions according to Einstein's diffusion equation:

$$r^2 = 6 \cdot D \cdot t \tag{2}$$

Where **t** is the time interval elapsed, **r** the average radius of distribution, and **D** the coefficient of diffusion, usually expressed as mm^2/s.

There are several ways to conduct a diffusion study. One is called Intra-Voxel Incoherent Motion – Diffusion Weighted Imaging (IVIM-DWI). It is a mathematical pseudo bi-exponential model that quantifies signal loss and restriction of the diffusion coefficient of free water molecules inside of tumor compared to those inside of healthy prostatic parenchyma. As a matter of fact, the curve fitting with the following equation takes two behaviors into consideration, slow and fast diffusion, related to cellularity and vascularization.

$$\frac{S}{S_0} = (1 - f) \cdot e^{-b \cdot D} + f \cdot e^{-b \cdot D^*} \tag{3}$$

Being $\mathbf{s/s_0}$ the normalized value of signal, **f** the vascular fraction, $\mathbf{D^*}$ the pseudo-perfusion coefficient. The vascular fraction **f** is a weighted parameter in regards to the proportion of vascular tissue in a voxel. The output of this equation is weighted average of the slow and fast diffusion, characterized and weighted by **D** and $\mathbf{D + D^*}$ respectively.

Though this model is rather physiologically appropriate, the distortion caused by the normalization of the signal reduces the signal-to-noise ratio and deforms the

original curve (Figure 4). Moreover, the biomarkers of this model, **D**, **D*** , and **f**, do not illustrate the correlation between pixels with the same behavior, and thus increase the uncertainty and produce results that are difficult to be interpreted.

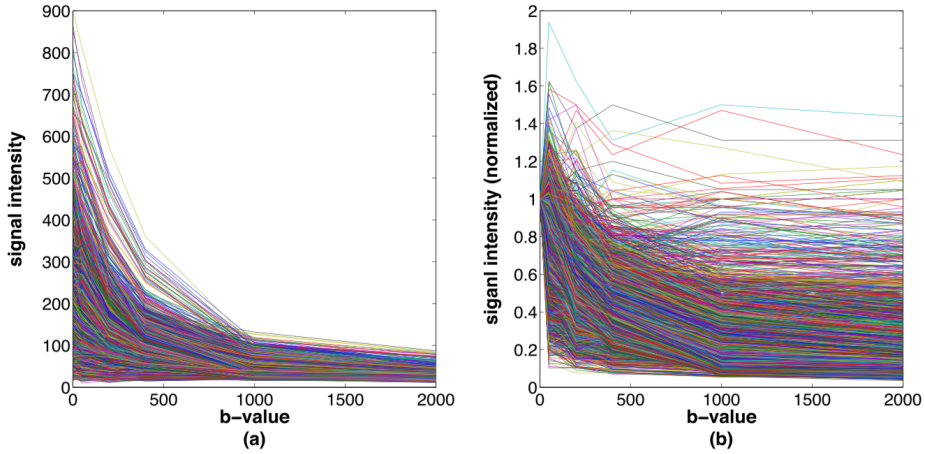


Figure 4. Signal attenuation in DW-MRI. (a) before normalization and (b) after normalization (Eric Aguado Sarrió and others, 2014).

Again, MCR-ALS can be used here so as to analyze the relation between pixels. By introducing a priori knowledge and other constraints, for example, non-negativity, unimodality, and shape constraints, the model is able to select non-orthogonal behaviors with more physiological meaning, and to model additive phenomena.

In the case of diffusion, a two-components model is applied. The biomarkers established in this model are **d1** (slow diffusion), **d2** (fast diffusion), and **RSS** (residual sum of squares). This model, whose biomarkers are complementary to those of the IVIM-DWI, has been proved capable of directly locating and grading the intensity of these behaviors in the DW-MRI images, improving the clinical diagnosis (Eric Aguado Sarrió and others, 2014).

The procedure of extracting the biomarkers will be more thoroughly described in the next chapter.

1.4. Image Coregistration

The coregistration of medical images of the same patient acquired in different modalities has been demonstrated beneficial in improving the prediction for pathological response and accuracy of locating regional lesions, typically in brains, breasts, lungs, etc (J. Ashburner and K. Friston, 1997; Yeong Yi An and others, 2015; Gerhard W. Goerres and others, 2002). Nowadays as the image post-processing techniques rapidly advance, more and more biomarkers are established and extracted from different modalities of images. That is why to accurately relate information from distinct images is of great interests and so widely investigated.

Alignment of a PET image with a MRI image was one of the earliest successful examples of coregistration, spawning a great variety of clinical applications. Its significance is that it permits structural or anatomical information to be deduced from images with less resolution, like PET images in this case. Normally, the image coregistration applications can be classified as:

- Multi-modality registration. The images are taken with different imaging technologies of the same part of the human body, extremely powerful when dealing with obtaining anatomical information from images with coarser resolution.

- Intra-modality registration. Images are taken over time with one modality in the same patient, which often happens when the goal is to monitor subtle changes or to inject different contrast media, for instance, MRI perfusion and diffusion sequences.

- Registration of images to physical space. Certain interventional or surgical treatments rely greatly on medical images to achieve productive outcome. Especially during treatment, the image-derived information must be aligned properly with actual physical space.

In this thesis one technique of image alignment (software Horos^{TM}) has been applied to synchronize the MRI perfusion and diffusion sequences, so that later the combined biomarkers can be analyzed using PLS-DA to determine whether the combination of these two modalities improves the predictability of prostate cancer.

Chapter 2. MATERIALS AND METHODS

2.1. Materials

Both the diffusion and perfusion sequences are performed along the volume of the prostate, at spatial planes corresponding to slices of the human body. All patient cases are from local hospitals, Valencia.

2.1.1. Perfusion sequences

Anonymized DCE-MR sequences of nine patients with surgically proven prostate cancer have been studied in this thesis. The images were taken with the contrast medium Gd-DOTA, under a resolution of 192 x 192 pixels, 47 non-equally spaced temporal samples, and 12 slices covering the whole pelvis area. As output, a 3D matrix of 564 DICOM images (12 slices x 47 time points) is constructed.

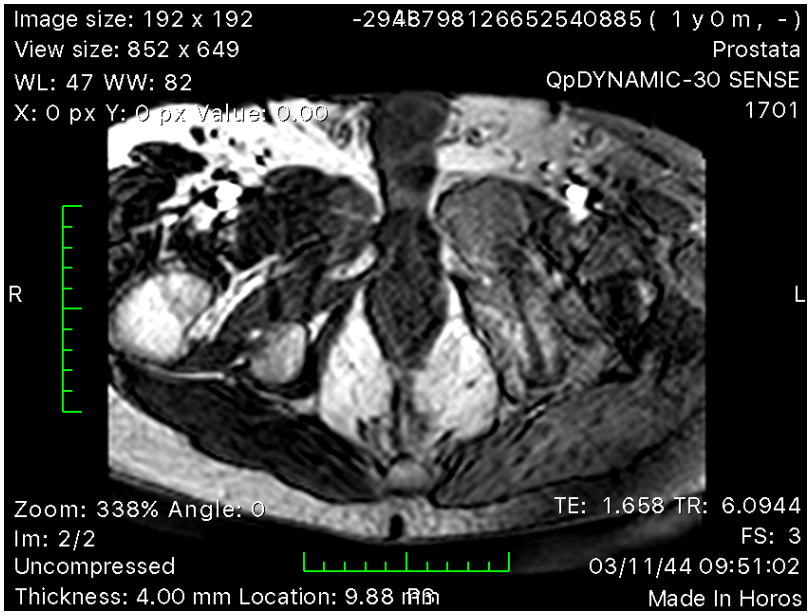


Figure 5. A dynamic-enhanced MRI image (perfusion) shown in Horos with DICOM format.

2.1.2. Diffusion sequences

The diffusion sequences were taken with DW-MRI of the same nine patients but at a different time, under a resolution of 192 x 192 pixels, also 12 slices including the whole pelvis area. Six different b-values (0, 50, 200, 400, 2000) s/mm^2 are administered. So the output is a 3D matrix of 72 DICOM images (12 slices x 6 time points) this time.

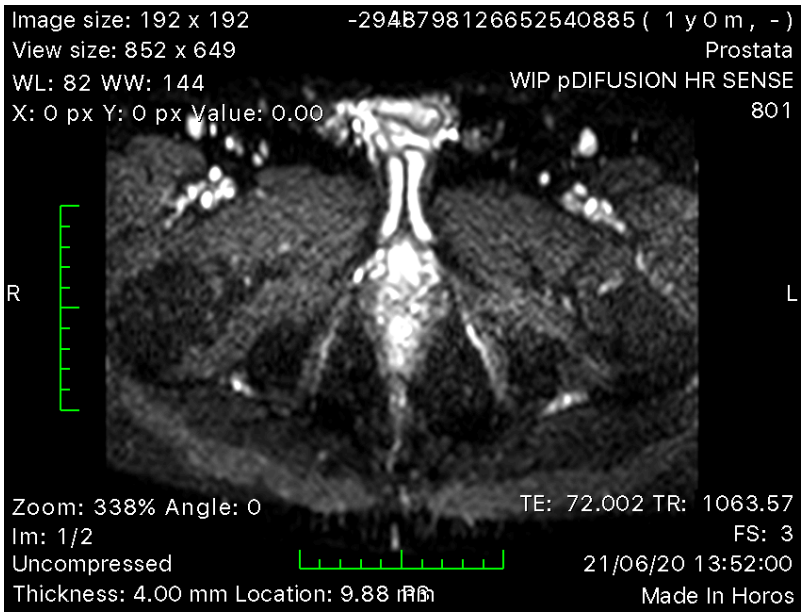


Figure 6. A diffusion weighted MRI image shown in Horos with DICOM format.

2.1.3. Software

The software applied for image coregistration is Horos (Purview, Annapolis, MD, USA) a free, open source medical image viewer based upon OsiriX as well as other open source libraries. OsiriX is a closed source program, whose FDA approved premium version costs 699 USD (suitable for medical use). Horos has been selected to conduct image alignment in this study because it is 64-bit, compared to the 32-bit free version of OsiriX (OsiriX Lite). This means that Horos conducts faster study loading than OsiriX Lite. Another important feature is that Horos is advertisement free while OsiriX Lite is not. Every time one opens OxiriX Lite, a window disclaiming “Not for Medical Use” appears, making the user experience less satisfactory to a certain extent.

```
================================================================================
Program:   OsiriX
 Copyright (c) OsiriX Team
 All rights reserved.
 Distributed under GNU - LGPL

 See http://www.osirix-viewer.com/copyright.html for details.
    This software is distributed WITHOUT ANY WARRANTY; without even
    the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR
    PURPOSE.
================================================================================*/
```

Figure 7. Tag appeared in codes of Horos.

The software later used for extraction of biomarkers and statistical analysis is Matrix Laboratory, commonly known as Matlab (The Mathworks Inc., Natick, MA, USA). The MCR-ALS and PLS-DA codes are based on codes previously programmed by Eric Aguado-Sarrió and José Manuel Prats-Montalbán, of the Multivariate Statistical Engineering Group, Universitat Politècnica de València, and later adapted and modified for this work.

2.2. Methods

2.2.1. Image Registration

All programmed codes for Horos is available at GitHub, where one can also share his/her own codes. Same as Osirix, Horos is written in the programming language Objective-C. Objective-C is a general-purpose, object-oriented and rather straightforward because of its Smalltalk-style messaging. For example, Figure 8 is part of the code DicomSeries, *@property* represents the categorical information contained in the tags of a dicom series. The rest of it amounts to some initial steps when opening dicom images, such as confirmation of filename for all images, identification of paths and key images, modification of series, etc.

```
71  @property(nonatomic, retain) NSNumber* stateText;
72  @property(nonatomic, retain) NSData* thumbnail;
73  @property(nonatomic, retain) NSNumber* windowLevel;
74  @property(nonatomic, retain) NSNumber* windowWidth;
75  @property(nonatomic, retain) NSNumber* xFlipped;
76  @property(nonatomic, retain) NSNumber* xOffset;
77  @property(nonatomic, retain) NSNumber* yFlipped;
78  @property(nonatomic, retain) NSNumber* yOffset;
79  @property(nonatomic, retain) NSSet* images;
80  @property(nonatomic, retain) DicomStudy* study;
81
82  - (NSSet*) paths;
83  - (NSSet*) keyImages;
84  - (NSArray*) sortedImages;
85  - (NSComparisonResult) compareName:(DicomSeries*)series;
86  - (NSNumber*) noFilesExcludingMultiFrames;
87  - (NSNumber*) rawNoFiles;
88  - (DicomSeries*) previousSeries;
89  - (DicomSeries*) nextSeries;
90  - (NSArray*) sortDescriptorsForImages;
91  - (NSString*) uniqueFilename;
92  @end
93
94  @interface DicomSeries (CoreDataGeneratedAccessors)
95
96  - (void) addImagesObject:(DicomImage *)value;
97  - (void) removeImagesObject:(DicomImage *)value;
98  - (void) addImages:(NSSet *)value;
99  - (void) removeImages:(NSSet *)value;
100
101 @end
```

Figure 8. Part of DicomSeries code of Horos.

The synchronization of images consists in aligning pixels that are congruent in two series, which means to identify the “counterpart” pixel of the one in the reference image, corresponding to the same anatomical point.

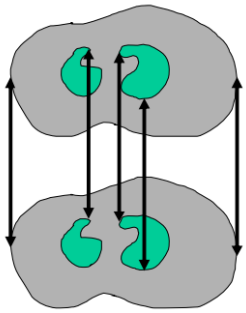


Figure 9. Image alignment at pixel level (Colin Studholme).

Due to the fact that the patients are scanned at different times, orientation, position and state of the target organ (prostate) would undoubtedly be distinct for perfusion and diffusion sequences. Besides, the patient location and orientation with respect to imaging system cannot be easily controlled or known. That is why a lineal combination of the two sequences does not suffice for a correct alignment at pixel level.

With Horos two series of images are synchronized automatically when loaded together at the workstation (Figure 10 (b)). In order to do that the function “sync series (same study)” under the “2D viewer” bar should be activated (Figure 10 (a)). The first series opened, always displayed at left, is the series taken as reference. This means that the secondly loaded series is moved upon the former to find “counterpart” pixels. In this study the perfusion sequences has been chosen as the reference set.

```
43  @class DICOMExport;
44  @class KBPopUpToolbarItem;
45
46  /** \brief  Window Controller for Orthogonal MPR */
47
48  typedef enum {SyncSeriesStateOff=0, SyncSeriesStateDisable=1, SyncSeriesStateEnable=2} SyncSeriesState;
49  typedef enum {SyncSeriesScopeAllSeries, SyncSeriesScopeSamePatient, SyncSeriesScopeSameStudy} SyncSeriesScope;
50  typedef enum {SyncSeriesBehaviorAbsolutePosWithSameStudy, SyncSeriesBehaviorRelativePos, SyncSeriesBehaviorAbsolutePos} SyncSeriesBehavior;
51
```

(a)

```
147  + (void) synchronizeViewer:(id)currentViewer;
148  + (void) synchronizeViewersPosition:(id) onlyViewerToBeSynchronized;
149  + (void) validateViewersSyncSeriesState;
```

(b)

Figure 10. Codes for synchronizing series. (a) Activation state of “sync series” function; (b) To synchronize only opened series.

Once loaded the two series, regions of interest (ROI) have been selected manually with the function “closed polygon” so as to fit the prostate as much as possible (Figure 11). Horos provides various tools for ROI selection in the pull-down menu, Ovals, Lines, Rectangles and Polygons. The Repulsor tool modifies the ROI created, by “pushing” the borders to make subtle changes. With the Propagate tool the ROI made in one image can be extended to the whole slice (47 for perfusion and 6 for diffusion). At last manual masks (ROIs) were made for each slice and same ROIs are used for both perfusion and diffusion sequences. This way the areas covered and position of ROIs are the same for both sequences to assure feasibility for later analysis (Figure 12).

```
2626        else
2627        {
2628                if( [[points lastObject] isNearToPoint: pt : scale/(thickness*backingScaleFactor) :[[curView curDCM] pixelRatio]] == NO)
2629                {
2630                        mypt = [[MyPoint alloc] initWithPoint: pt];
2631
2632                        [points addObject: mypt];
2633                        [mypt release];
2634
2635 //                     NSLog(@" [ROI, mouseRoiDown] adding point for polygon...");
2636 //                     NSLog(@" [ROI, mouseRoiDown] slice : %d", slice);
2637                        [zPositions addObject:[NSNumber numberWithInt:slice]];
2638
2639                        clickPoint = pt;
2640                }
2641                else    // Click on same point as last object -> STOP drawing
2642                {
2643                        mode = ROI_selected;
2644                }
```

Figure 11. Codes for Polygon selection.

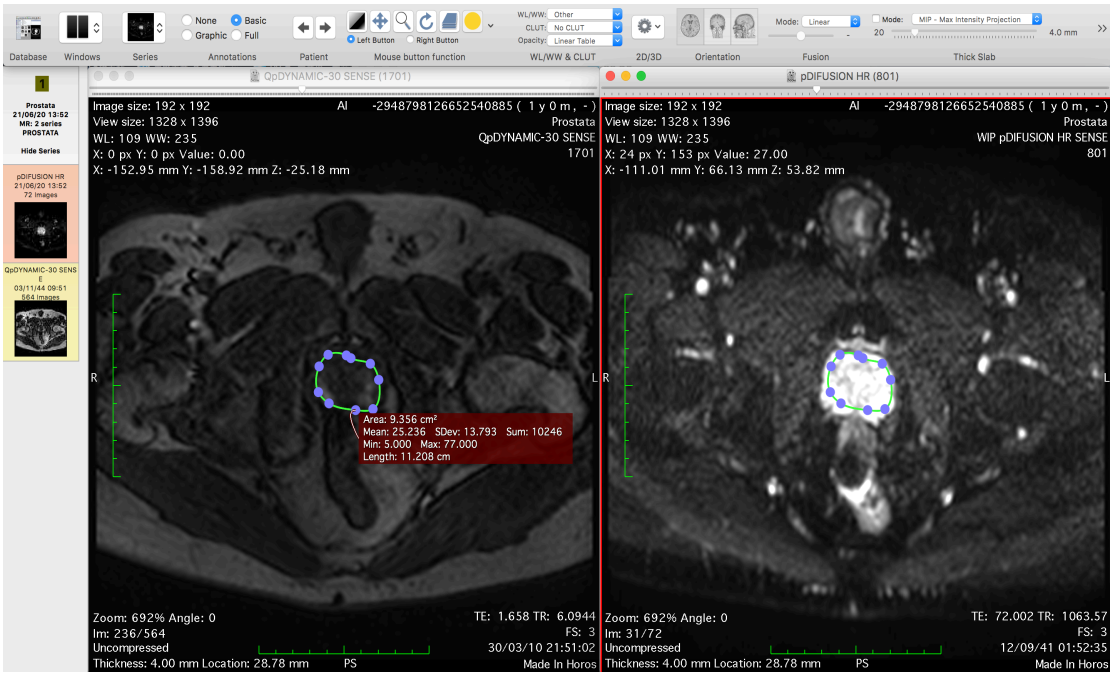


Figure 12. Workstation of one case (out of nine) with ROIs selected (shown as green).

As represented in Figure 12, prostate is normally located in the middle of the image, possessing a more bright contrast. Yet not all that bright zone was selected as a precaution against disturbance towards the behaviors and introduction of more noise into PLS-DA model. Next, for the purpose of reducing computational cost as well as noise, pixels outside of ROIs were set to zero. Finally the new constructed series were exported to DICOM images to be analyzed with MATLAB.

2.2.2. Extraction of biomarkers

The exported DICOM images, like the originals, can be considered as a 3D matrix data, which needs to be unfolded into a 2D matrix for MCR model. Because the resolution is the same for perfusion and diffusion sequences, each slice of the case is unfolded as a 2D matrix of size 36,864 x time points/slice, being 6 for diffusion and 47 for perfusion. Then each slice is stacked one below the other in rows forming a second 2D matrix of size 442,368 x time points/slice. This way the wanted information is kept intact as the defined behaviors are forced to maintain the same internal correlation structure.

- **MCR-ALS for perfusion**

  According to radiologists and professionals working with these images, the first, second, ninth, and tenth slice does not include prostate area. Therefore these slices were discarded later during the process.

  As previously described in chapter 1, there are three dynamic behaviors in the perfusion model, type A, NT and VT. MCR is an iterative method that performs a bilinear decomposition of matrix **X** by means of an ALS optimization (José Manuel Prats Montalbán and others, 2013):

$$\mathbf{X} = \mathbf{C}\mathbf{S}^T + \mathbf{E} \quad (4)$$

  Each dynamic behavior is modeled in rows of the matrix $\mathbf{S}^T$, the relative importance of each behavior corresponding to each pixel is contained in the matrix **C**, and **E** is the residue matrix.

  As it was mentioned above, *a priori* knowledge is introduced in the MCR-ALS model and can be employed as an initial estimation. Apart from that, it helps to

select appropriate constraints so that the iterative model can yield finite and interpretable solutions.

To know what is to be expected is very helpful to the correct definition of the "pure" dynamic behaviors existing in the images. After the initial estimation, principal component analysis (PCA) can be applied to validate the *a priori* knowledge and to determine the number of the behaviors. To do so the umber of latent variables, or the principal components (PCs), with the highest variance are counted. Once the number is determined, the purest dynamic behaviors can be sought by using Simple-to-use Interactive Self-modeling Mixture Analysis (SIMPLISMA) based algorithm of the software available at the MCR homepage (MCR homepage, 2016). In this case, three dynamic behaviors have been identified:

- Type AIF: drastic enhancement, corresponding to AIF.
- Type NT: low progressive enhancement with no decay, corresponding to healthy tissue.
- Type VT: delayed drastic enhancement with widened and slow decay, corresponding to highly vascularized tissue (tumor).

The next move would be to find the dynamic behaviors present in each pixel, which could be from zero to all of them. And because the goal here is to "decipher" biological information out of the database, constraints like non-negativity is employed for the model as a pixel with negative intensity does not make any physiological sense. Even though *a priori* information is applied for determination of dynamic behaviors in pixels, it would be better to leave it out when applying local MCR models for each case. This way possible bias resulted from *a priori* knowledge is prevented and the attained AIF is best fitted in each case.

After applying non-negativity constraints on pixel intensity values and on the dynamic behaviors, the **C** matrix can now be folded back into the original size (192 x 192) to obtain the distribution maps. Only through these maps can the pixels with more correlation to the dynamic behaviors be located. However, the local MCR model discards the type AIF behavior, since there are no arteries inside the prostate. The newly emerged behavior is identified by doctors as the non-physiological contrast media arrival (CMA), or simply the artificial effect, which seems to affect the VT and NT behaviors to certain degree (Figure 13).

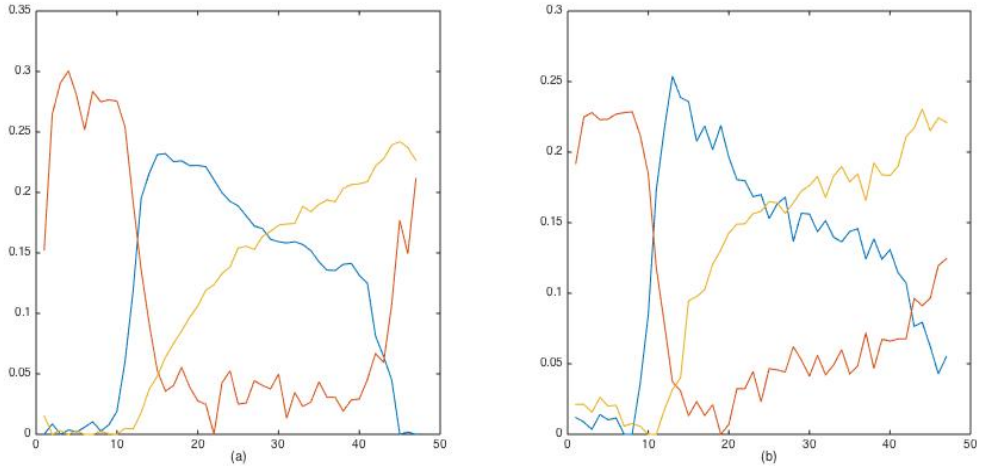


Figure 13. (a) Initially estimated dynamic behaviors, type NT (yellow), type VT (blue), type contrast media arrival (red). (b) Dynamic behaviors patterns of one case.

Once all nine cases have been examined to show similar patterns as Figure 13(a) (pixels with extreme value would deform the patterns and catastrophically influence the PLS-DA model), the same ROIs defined by doctors for the pharmacokinetic model have been used for segmentation. The ROIs are the regions of interest marked for each case, which are composed of either tumor or the healthy tissue of prostate, denominated as DL - dominant lesion or HP – healthy peripheral. During segmentation, the ROIs serve as sample zones used to relate regions in the image with comparable biological composition-related properties.

◆ MCR-ALS for diffusion

With diffusion the model does not vary much from the basic equation, (Equation 4). This time the matrix **X** contains the signal spectrum for each pixel in rows; $\mathbf{S}^T$ contains in rows the diffusion behaviors modeled; **C** contains the relative contribution of modeled behavior for each pixel, and **E** is the residue matrix.

The studied phenomena in diffusion are the slow diffusion (d1), associated to cellularization, and fast diffusion (d2), associated to vascularization. Assuming that the signal spectrum in a pixel can be expressed as a weighted sum of different decreasing exponential functions, so the following model is proposed in previous investigation (Eric Aguado Sarrió, 2014):

$$s_j = \sum_{i=1}^{I} c_{ij}\left(\alpha_i e^{-\beta_i b}\right); \quad \alpha_i, \beta_i, c_{ij} \geq 0 \tag{5}$$

where **I** means the number of exponential functions used, the rest are exponential parameters determined sequentially. Details can be found in (Eric Aguado Sarrió, 2014).

Constraints need to be employed here as well. As mentioned earlier in chapter 1, non-negativity constraints are used in matrices **S** and **C**, for the behaviors and their relative contribution lied in a pixel have to be positive to make sense. Unimodality constraints are only used for the matrix **S**, where the modeled behaviors are monotonically decreasing (Eric Aguado Sarrió and others, 2014). At last shape constraints are applied also to matrix **S** inside ALS to assure that the modeled behaviors are expressed as exponential decay.

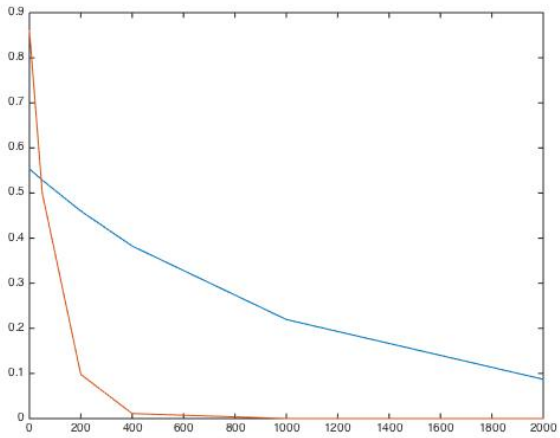


Figure 14. The exponential decay of the modeled behaviors.

Blue: slow diffusion (d1); red: fast diffusion (d2).

The $\mathbf{S}^{\mathbf{T}}$ and **C** matrices gather at each pixel position the behaviors as well as their relative contribution. Like in perfusion, the matrix **C** is then folded into original spatial dimension (192 x 192) to build distribution maps. The idea behind it is to relocate the pixels more related to the corresponding behaviors in each column of **C**, because these maps generated from **C** are the ones performing as imaging biomarkers. Later the distribution maps of the RSS can be used to validate the result, as the pixels not well fitted by the model are included in it.

The next step is to apply the same ROIs used for perfusion to diffusion models segmentation. It must be stressed that the ROIs are applied with perfusion masks because the perfusion sequences were taken as reference during image

alignment.

With the MCR model finalized, the $\mathbf{X_{DIF}}$ matrix containing the three biomarkers (d1, d2, RSS) in columns is obtained. The rows are all the pixels "filtered" by ROIs. Equally, the matrix $\mathbf{X_{PER}}$ is constructed with the three biomarkers NT, VT, and RSS of perfusion. The AIF behavior is taken out of the matrix for improved precision. And a matrix **Y** contains the information about classification of each pixel, whether it belongs to class DL or HP. Finally a new **X** matrix including all six biomarkers is formed, completing the combination, which is easily done with Matlab (Figure 15).

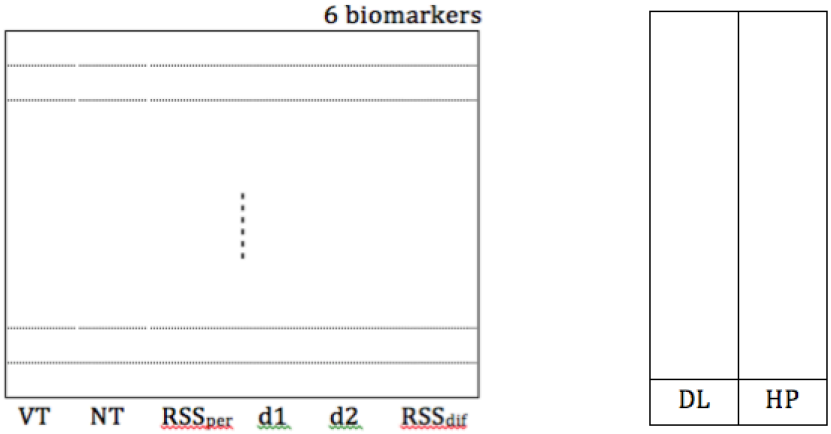


Figure 15. The X and Y matrices for PLS model. (a) X matrix; (b) Y matrix.

2.2.3. PLS-DA analysis

The Partial Least Squares regression is one of the possible models of prediction that tries to find the fundamental relations between two data matrices, **X** (studied images) and **Y** (output data), maximizing the covariance of **X** and **Y** by a linear multivariate model. That is to say, PLS conducts a projection of the matrix **X** to approximate it properly and to maximize the correlation with matrix **Y**. The output data can be formed by other images or simply data of different nature, such as analytical or mechanic parameters, temporal evolution, and even artificial dummy variables associated with each type of the image in the case of a discriminant analysis. This way the new predicted output data of new images can be calculated through constructed PLS model.

When it comes to multivariate images, it is also denominated as multivariate image regression, the MIR. For instance, for cases aiming at tumor or lesion detection, quantification and localization inside of a patient, it is often referred to an inferential prediction model, whose result is an output image. However, a Partial Least Squares Discriminant Analysis (PLS-DA) has been chosen in this work, as the matrix Y is categorical, created with *dummy* variables (ones and zeros). Generally, the PLS-DA is applied for classification of new projected images. The Figure 16 shows the basic idea for PLS-DA model construction.

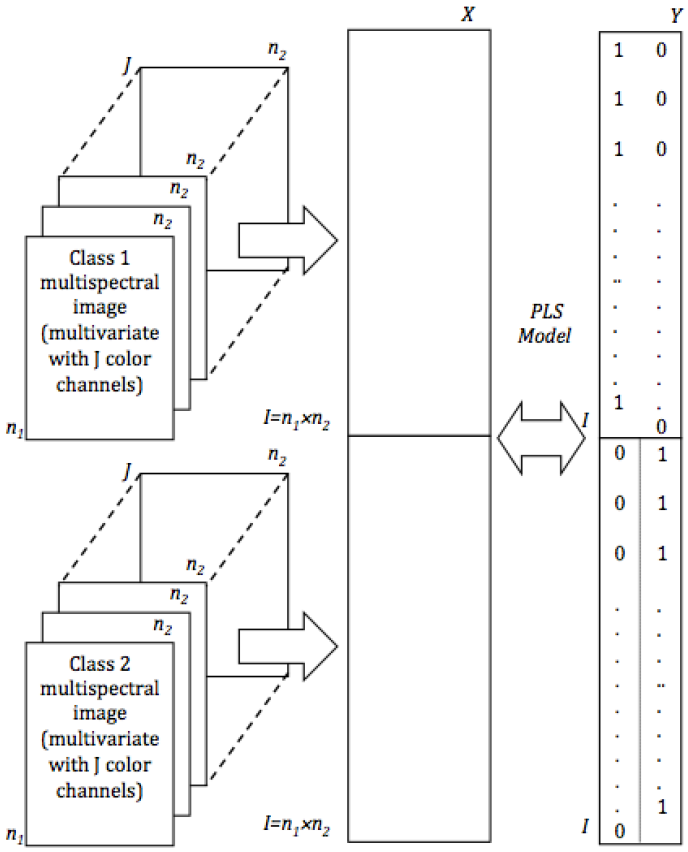


Figure 16. Framework of a PLS-DA model construction (José Manuel Prats Montalbán, 2005).

Each of the images analyzed is placed one below another, creating *dummy* variables for each of the pattern images. Normally the PLS-DA is calculated with predictive algorithms (a PLS model) on unfolded data structure:

$$\boldsymbol{X_{ij}} = \sum_{j=1}^{F} \boldsymbol{t_{if}}\, \boldsymbol{p_{if}} + \boldsymbol{r_{ij}(x)} \tag{6}$$

$$\boldsymbol{Y_{ij}} = \sum_{j=1}^{F} \boldsymbol{u_{if}}\, \boldsymbol{q_{if}} + \boldsymbol{r_{ij}(y)} = \sum_{j=1}^{F} \boldsymbol{t_{if}}\, \boldsymbol{q^{*}_{if}} + \boldsymbol{g_{ij}} \tag{7}$$

where **t**, **u** are vectors scores of the matrix **T** and **U** in the **X** and **Y** plane respectively; **r** is the sum of squares of the residue; **g** is the global residue; **q** are the weights that combine the variables of **Y** to form the scores **u** so that the covariance between **Y** and **X** is maximized.

Given that:

$$u_{if} = t_{if} b_{\square} + h_{if} \tag{8}$$

$$\widehat{u_{if}} = t_{if} b_f \tag{9}$$

leaving the defined b_f as

$$b_f = {u_f^T t_f}\Big/{t_f^T t_f} \tag{10}$$

If expressed with matrices:

$$\boldsymbol{X = TP^T + R_x} \tag{11}$$

$$\boldsymbol{Y = UQ^T + R_Y = TQ^{*T} + G} \tag{12}$$

where **R** stands for external relations and **U** the internal relations (matrix of scores in the **Y** plane); **T** is the matrix of scores **t** in the **X** plane; **Q** itself is the matrix of correlation between **t(X)** and **Y**, related to the score u; **P** is the direction of the component in **X**, commonly called as loadings.

The fundamental reason for choosing a PLS is its operation as a prediction model, from which a standard regression equation for prediction is applied:

$$\boldsymbol{Y_{pred} = XB_{PLS}} \tag{13}$$

where **B** is the calculated regression coefficient. $\mathbf{Y_{pred}}$ is a vector with precise information in each variable m in the matrix **Y**, weighted by the calculated coefficiente $\mathbf{b_m}$.

$$y_m = \square_{1m} x_1 + b_{2m} x_2 + \cdots + b_{km} x_k + f_m \tag{14}$$

And because:

$$\boldsymbol{T = XW^* = XW(P^TW)^{-1}} \tag{15}$$

$$\boldsymbol{Y_{pred} = XB_{PLS} = TQ^T = XW(P^TW)^{-1}Q^T} \tag{16}$$

Being **W** the matrix of correlation between **X** and **u(Y)**, and $\mathbf{W}^*$ the matrix of correlation between matrices **X** and **Y.**

Therefore:

$$\boldsymbol{B_{PLS} = W(P^T W)^{-1} Q^T} \quad (17)$$

A PLS model facilitates the interpretation as well as the explication of the results to non-experts, when the number of the variables is tremendous. Algorithm SIMPLS (José Manuel Prats Montalbán, 2005) is used here to calculate PLS, because it is much faster especially when matrix **X** is of large size.

2.2.4. Leave one out Cross-Validation

After the PLS model, a validation method has been applied to verify the results. Particularly, the cross-validation evaluates the predictive ability of a model by calculating the "right" number of components to retain in the model. Usually it is used in prediction realm to assess how well a model performs in practice, when the number of observations is not very high.

Normally a predictive model is given a training dataset of *known data* to run as the first step, and a test dataset of *unknown data* to validate the outcome of the training data. Typically 30 percent of the data is divided and put into the test data group. This is one of the most classical methods for validation, the training set versus the test set. It is a simple and straightforward tool using independent unknown dataset. For instance, the predictive accuracy in this case is often measured by the mean squared error (MSE) on the test set. However, if the available data were rather little, or the distribution of the data is no good, not only would this method waste data, but also would there be greater probability for the performance to have high variance. When that happens, there are really no other choices but to consider the validation unreliable for model performance.

Therefore the leave one out cross-validation method (LOOCV) is often referred to as an alternative. It is somewhat similar but more sophisticated than the original training/test version. Here the accuracy is also measured by the average of the error (MSE) of all **n** observations, $\mathbf{y_1}$, … $\mathbf{y_n}$. The procedure of this method is summarized as follows:

i. Leave out the observation **i** from the data set. Then fit the model with the rest of the data except observation **i**.

ii. Test the accuracy of the model with the observation **i**, computing the error $(e_i^* = y_i - \hat{y}_i)$ (predictive residual) for the omitted observation.

iii. Repeat the process with all data (**n** observations) in the set and calculate the accuracy for each one.

iv. The LOO accuracy now amounts to the average of the accuracy of **n** observations, computing the MSE from $e_1^*, \ldots, e_n^*$.

This way the available data is more efficiently used and problems like over-fitting can be limited, by determining the appropriate number of components for the model. Plus it provides insight on how the model performs while generalizing to independent dataset. It is also convenient as the human bias is excluded from the model with the same database. To sum up, cross validation combines averages measures of prediction error to correct for the training error and thus derives a more accurate estimation of model prediction performance (Robert Grossman and others, 2010).

In this thesis the LOOCV is selected due to the size of data set, 190 observations (pixels), and because it is less time-consuming.

Chapter 3. RESULTS

Results

MCR analysis was performed on all nine cases, obtaining the relative importance of the six biomarkers respectively for the MCR local models. Figures 17 and 18 display the local distribution maps of the behaviors in perfusion and in diffusion.

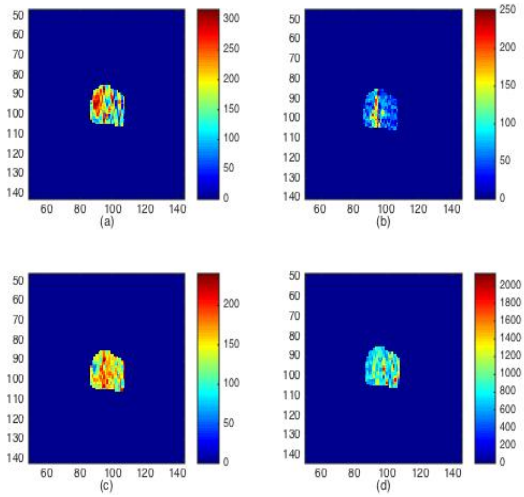


Figure 17. Local distribution maps of the dynamic behaviors in perfusion:

(a) type VT, (b) type CMA, (c) type NT, (d) RSS.

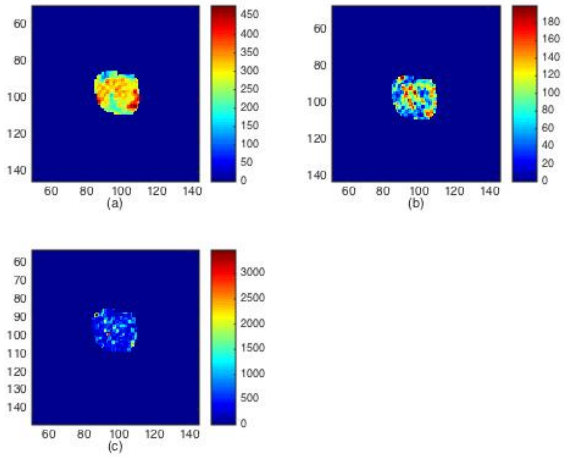


Figure 18. Local distribution maps of behaviors in diffusion:

(a) slow diffusion (d1), (b) fast diffusion (d2), (c) RSS

For the PLS model to work correctly, the matrices **X** and **Y** need to be well balanced before the maximization of the covariance. Accordingly, by observing the histograms of components in the matrix **X**, filters have been applied to omit pixels whose value is larger than 500. During the prediction, or more precisely the classification of each pixel, if one pixel

is categorized as healthy, it can either be TN (true negative) or FN (false negative). Likewise, for a pixel identified as tumor, it could be a TP (true positive) or FP (false positive). With these basic parameters, two distinct factors can later be deduced:

$$Precision = \frac{TP}{TP+FP} \quad Recall = \frac{TP}{TP+FN} \tag{18}$$

These two are used to calculate the f-score, which serves as the index for the goodness of prediction of the model. The closer is the f-score to 1 (maximum value), the better predicts the model.

$$fscore = \frac{2 \cdot precision \cdot recall}{precision + recall} \tag{19}$$

As results, the PLS model shows a goodness of prediction 0.8144, with TN 0.3947, TP 0.4158, FN 0.0526, and FP 0.1369, accuracy 0.8105. At this point the model has performed an impressively good prediction. The following picture (Figure 19) shows the original classification (matrix **Y**) on the left and the prediction by the model (Y_{pred}) on the right, represented also with dummy variables (1s and 0s), by setting the larger probabilities to 1 and leaving the minors to 0.

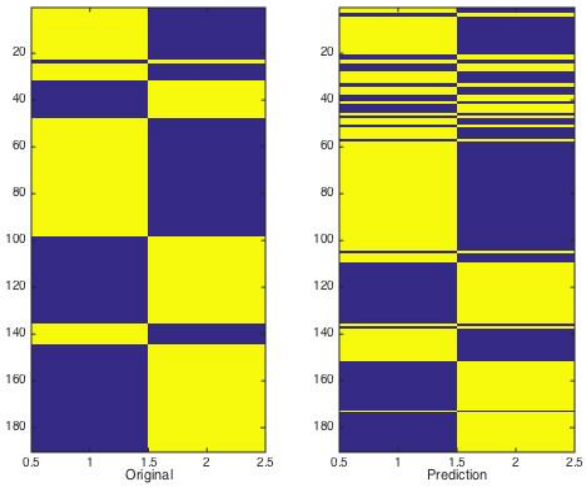


Figure 19. Results of PLS-DA model, with matrix Y on the left and Y_{pred} on the right.

PLS model has also been applied with biomarkers of perfusion and diffusion separately. For diffusion, the f-score only reaches 0.4895, with TN 0.4316, TP 0.1842, FN 0.2842, FP 0.1.

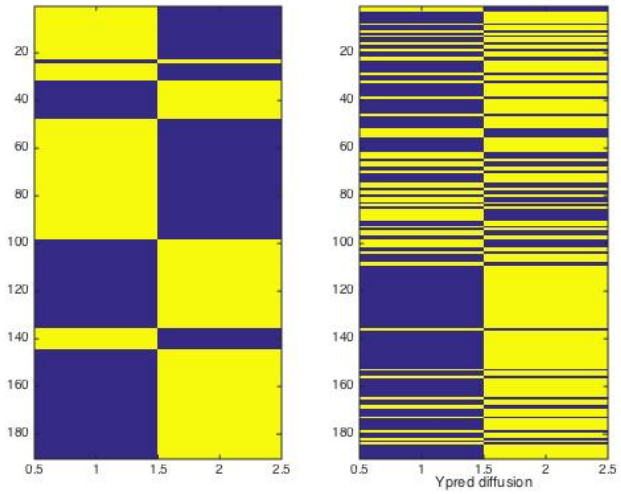


Figure 20. PLS prediction with only diffusion biomarkers (results on the right).

For perfusion alone, the results are also quite frustrating. Calculated f-score possesses a value of 0.4113, with TN 0.4105, TP 0.1526, FN 0.3158, FP 0.1211. In plain sight, the combined biomarkers perform a prediction a lot better than those of each sequence alone.

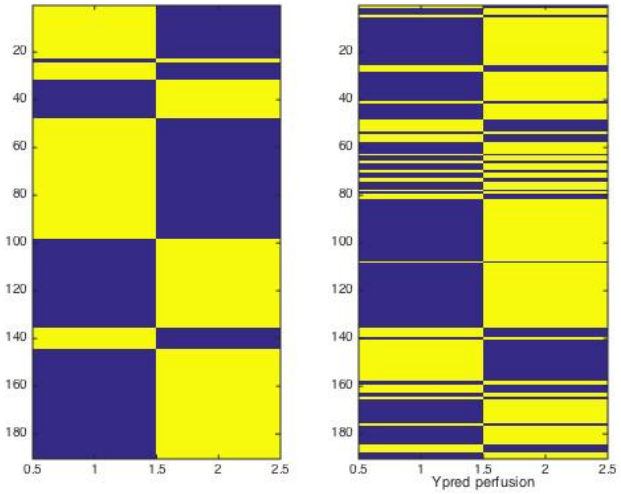
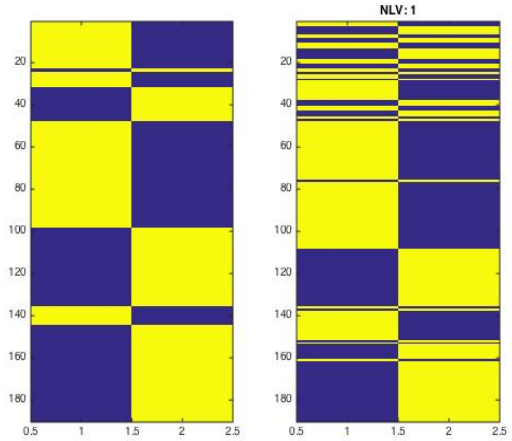


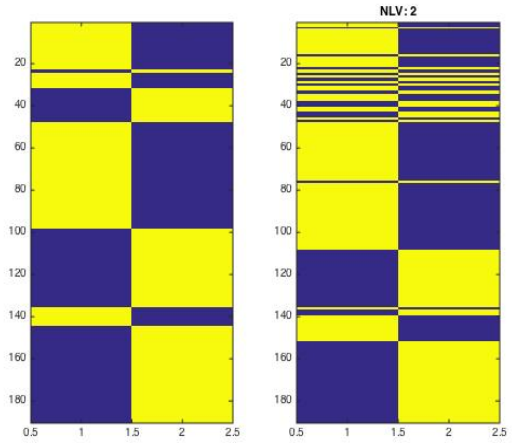
Figure 21. PLS prediction with only perfusion biomarkers (results on the right).

For validation purposes, leave one out cross validation have been conducted to verify the results of the PLS-DA model. The initial number of latent variable has been set for six. And the dataset (matrix **X**) has been managed with auto-scaling to achieve better data distribution. Auto-scaling is a data processing technique in which each column is subtracted by average and then divided by typical deviation. The underlying idea is for each biomarker to have the same importance, or rather, the same contribution to the model. The LOOCV shows a two-component proved model, for there is no significant improved prediction performance starting from three latent variables. This means that if one goes further introducing more than two components into the model, it would be over-fitted and yield unreliable results, no matter how good the accuracy seems.

This two-component model gives an error of 37 wrong observations out of 190 in total, hence an accuracy of 0.8053. The goodness of prediction (f-score) is also slightly decreased, with a value of 0.8122, (TN 0.3842, TP 0.4211, FN 0.0474, FP 0.1474).



(a)



(b)

Figure 22. The results of LOOCV of one-component (a) and two-component (b).

Obvious improvement in the two-component model can be observed in the image.

CHAPTER 4. CONCLUSIONS

Conclusions

In this work, the software Horos has been applied for image alignment instead of using the classical mathematic method. The automatic synchronization of the series makes it a rather simple and rapid alternative. It ought to be noted that Horos is not FDA approved while Osirix-MD is. This main difference makes Horos inappropriate for clinical diagnosis. However, as Horos is based upon Osirix, it is presumable that same outcome can be achieved using Osirix-MD.

Expert knowledge is preferred during the selection of ROIs with Horos, since it is of great importance to select all prostate avoiding introducing too much noise. Excessive noise would deform the dynamic behaviors and thusly affects the biomarkers extraction. Along the present work, the selected ROIs were restricted to a rather small area because of this very reason. Consequentially, some of the "good" pixels were left out, reducing the number of observations. That is why in the future it is recommendable to study the model with a larger data set and maybe with other different validation methods to assess the model performance.

The obtained results show great potential in further applications, particularly the prostate cancer diagnosis on a patient basis as a fast method to look into both tissue cellularization and vascularization. The combined biomarkers of perfusion and diffusion could also give a more comprehensive insight into tumor stages, its characteristics, etc. As a preliminary study, the next step could be to investigate and to evaluate with more depth the precision and differences between the alignment by Horos and by numerical methods. Besides, another interesting question that remains to be answered is whether the combination or addition of those widely applied biomarkers (the IVIM and pharmacokinetic model), which are more theoretically physiological, would improve prediction performance.

BUDGET

Budget

The budget of this work mainly comes from personnel expenses. I have been worked on this project for nearly 300 hours. Should I, as an engineer, charge 15€/h, the accrued cost would then be 4500€ in total. José Manuel Montalbán and Eric Aguado Sarrió should be considered as consultants, at a cost of 30€/h. And they have worked with me for 50h and 100h respectively.

The software used were Horos, student version of Matlab and Microsoft Office. Horos is a free software, while Matlab and Microsoft Office are purchased student version. Same as Osirix, Horos can only be used with operating system OS X. Therefore a Macbook Pro (15 inches) is applied in this work. All programmed codes were properties of the Multivariate Statistical Engineering Group, Universitat Politècnica de València, Valencia, Spain.

Description	Unit Price	Amount	Total
Engineer	15€/h	300h	4.500 €
Consultant 1	30€/h	50h	1.500 €
Consultant 2	30€/h	100h	3.000 €
Materials	0 €	9 cases	0 €
Matlab	69€/unit	1	69 €
Microsoft Office*	149€/unit	1	149 €
Horos	0 €	1	0 €
Apple Macbook Pro	2249€/unit	1	2.249 €
Accessory*	89€/unit	1	89 €
Total			11.556 €

Table 1. Budget Table. * The student version of Microsoft Office refers to Office Home and Student 2016 for Mac; Accessory is Magic Mouse 2 by Apple Inc. Cupertino, California, U.S.

If the proposed methodology were to be used by a company for clinical use in the United States, other than personnel expenses, Osirix-MD will cost 699$ (627.53€); the price of Matlab will hinge on selected toolboxes and number of employees using Matlab at the same time. Usually the Network Named User License can be purchased in order to buy fewer licenses than the headcount of the company.

References

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, et al. Cancer statistics 2008. CA Cancer J Clin. 2008; 58:71---96.
2. Miller DC, Hafez KS, Stewart A, Montie JE, Wei JT (September 2003). "Prostate carcinoma presentation, diagnosis, and staging: an update form the National Cancer Data Base". Cancer 98 (6): 1169–78.
3. Hankey BF, Feuer EJ, Clegg LX, Hayes RB, Legler JM, Prorok PC, Ries LA, Merrill RM, Kaplan RS (June 16, 1999). "Cancer surveillance series: interpreting trends in prostate cancer—part I: Evidence of the effects of screening in recent prostate cancer incidence, mortality, and survival rates". J Natl Cancer Inst 91 (12): 1017–24.
4. David Martí Aguado, Los biomarcadores de imagen farmacocinética permiten diagnosticar y localizar el carcinoma prostático con resonancia magnética, 2014
5. World Cancer Report 2014. World Health Organization. 2014. pp. Chapter 5.11. ISBN 9283204298
6. Hsing AW, Chokkalingam AP (2006). "Prostate cancer epidemiology". Frontiers in Bioscience 11: 1388–413.
7. Beuzeboc P, Soulié M, Richaud P, Salomon L, Staerman F, Peyromaure M, Mongiat-Artus P, Cornud F, Paparel P, Davin JL, Molinié V (December 2009). Fusion genes and prostate cancer. From discovery to prognosis and therapeutic perspectives. Prog. Urol. (in French) 19 (11): 819–24.
8. Andriole GL, Crawford ED, Grubb III RL, Buys SS, Chia D, Church TR, et al. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med. 2009; 360:1310---9.
9. Schröder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. N Engl J Med. 2009; 360:1320---8.
10. José Manuel Prats-Montalbán, Roberto Sanz-Requenab, Luis Martí-Bonmatí and Alberto Ferrer, Prostate functional magnetic resonance image analysis using multivariate curve resolution methods
11. Jackson ASN, ReinsbergSA, SohaibSA, Charles-EdwardsEM, JhavarS, Christmas TJ, Thompson AC, Bailey MJ, Corbishley CM, Fisher C, Leach MO, Dearnaley DP. Dynamic contrast-enhanced MRI for prostate cancer localization. Br. J. Radiol. 2009; 82: 148–156.
12. Leach MO, Brindle KM, Evelhoch JL, Grif ths JR, Horsman MR, Jackson A, Jayson GC, Judson IR, Knopp MV, Maxwell RJ, McIntyre D, Padhani AR, Price P, Rathbone R,

Rustin GJ, Tofts PS, Tozer GM, Vennart W, Waterton JC, Williams SR, Workman P. The assessment of antiangiogenic and antivascular therapies in early-stage clinical trials using magnetic resonance imaging: issues and recommenda- tions. Br. J. Cancer 2005; 92: 1599–1610.

13. Barrett T, Brechbiel M, Bernardo M, Choyke PL. MRI of tumor angiogenesis. J Magn Reson Imaging. 2007 Aug; 26(2): 235-49.
14. Juan A, Maeder M, Hancewicz T, Duponchel L, Tauler R. Chemometric tools for image analysis. In Infrared and Raman Spectro- scopic Imaging, Wiley-VCH Verlag GmbH & Co. KGaA, 2009; 65–109.
15. E. Aguado-Sarrió, J.M. Prats-Montalbán, R. Sanz-Requena, L. Martí-Bonmatí, A. Alberich-Bayarri, A. Ferrer, Prostate Diffusion Weighted-Magnetic Resonance Image analysis using Multivariate Curve Resolution methods
16. L. Martí-Bonmatí, A. Alberich, R. Sanz, J. Sánchez. State of art on living images-MR Diffusion/Perfusion. GE Healthcare. 2010.
17. D. Le Bihan, E. Breton, D. Lallemand, M. L. Aubin, J. Vignaud, y M. Laval-Jeantet. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. Radiology, 497-505, 1988.
18. Quint LE, Van Erp JS, Bland PH. Carcinoma of the prostate: MR images obtained with body coils do not accurately reflect tumor volume. AJR 1991; 156:511–516
19. Tofts PS, Brix G, Buckley DL, Evelhoch JL, Henderson E, Knopp MV, Larsson HB, Lee TY, Mayr NA, Parker GJ, Port RE, Taylor J, Weisskoff RM. Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: standardized quantities and symbols. J. Magn. Reson. Imaging 1999; 10: 223–232.
20. Multivariate curve resolution homepage, http://www.mcrals.info/.
21. J. Ashburner and K. Friston, Multimodal Image Coregistration and Partitioning—A Unified Framework, NEUROIMAGE 6, 209–217 (1997)
22. Yeong Yi An, Sung Hun Kim, Bong Joo Kang, and Ah Won Lee, Treatment Response Evaluation of Breast Cancer after Neoadjuvant Chemotherapy and Usefulness of the Imaging Parameters of MRI and PET/CT, J Korean Med Sci 2015; 30: 808-815
23. Gerhard W. Goerres, MD1; Ehab Kamel, MD1; Burkhardt Seifert, PhD2; Cyrill Burger, PhD1; Alfred Buck, MD1; Thomas F. Hany, MD1; and Gustav K. von Schulthess, MD, PhD1, Accuracy of Image Coregistration of Pulmonary Lesions in Patients with Non–Small Cell Lung Cancer Using an Integrated PET/CT System, THE JOURNAL OF NUCLEAR MEDICINE • Vol. 43 • No. 11 • November 2002
24. David J Hawkes, Derek LG Hill, Lucy Hallpike and Dale L Bailey, Coregistration of Structural and Functional Images*
25. Horos homepage, https://www.horosproject.org/
26. Colin Studholme, Image Registration in Medical Imaging, Department of Radiology and Biomedical Imaging University of California San Francisco UCSF/UCB Joint

Graduate Group in Biomedical Engineering

27. José Manuel Prats Montalbán, Control Estadístico de Procesos mediante Análisis Multivariante de Imágenes, 2005
28. Svante Wold, Michael Sjo¨stro¨m, Lennart Eriksson, PLS-regression: a basic tool of chemometrics, Chemometrics and Intelligent Laboratory Systems 58 Ž2001. 109–130
29. Rob J Hyndman, Why every statistician should know about cross validation, october, 2010. http://robjhyndman.com/hyndsight/crossvalidation/
30. Grossman, Robert; Seni, Giovanni; Elder, John; Agarwal, Nitin; Liu, Huan (2010). Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions. Morgan & Claypool.
31. GitHub - horosproject – horos page. https://github.com/horosproject/horos