Multi modal Medical Registrati on and Fusion. Applicati on to Endometriosis Surgery

Youssra El Bennioui

RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

Université de Toulouse

Doctorat de l'Université de Toulouse

préparé à Toulouse INP

Fusion et recalage d'images médicales multimodales. Application à la chirurgie de l'endométriose

---

Thèse présentée et soutenue, le 8 novembre 2024 par

Youssra EL BENNIOUI

École doctorale
EDMITT - Ecole Doctorale Mathématiques, Informatique et Télécommunications de Toulouse

Spécialité
Informatique et Télécommunications

Unité de recherche
IRIT - Institut de Recherche en Informatique de Toulouse

Thèse dirigée par
Jean-Yves TOURNERET et Adrian BASARAB

Composition du jury
Mme Mireille GARREAU, Présidente, Université de Rennes 1
M. Jean-Louis DILLENSEGER, Rapporteur, Université de Rennes
Mme Su RUAN, Rapporteuse, Université de Rouen
M. Abderrahim HALIMI, Examinateur, Heriot-Watt University
Mme Oumaima EL MANSOURI, Examinatrice, Abelio
M. Julian TACHELLA, Examinateur, CNRS Rhône Auvergne
M. Jean-Yves TOURNERET, Directeur de thèse, Toulouse INP
M. Adrian BASARAB, Co-directeur de thèse, Université Claude Bernard Lyon 1

Acknowledgments

Writing these ackowledgments marks the conclusion of more than three years of hard work, at times exhausting, but ultimately so rewarding. A PhD journey is far from a smooth ride, thankfully, many people have accompanied and supported me along the way, helping me steer my ship safely to shore. These few lines are dedicated to them.

I would like to begin by expressing my deepest gratitude to my supervisors, Professors Jean-Yves TOURNERET and Adrian BASARAB. Your guidance, significant support, and encouragement throughout this journey have been invaluable. Over these last three years, your constructive feedback, brilliant ideas, along with your kindness and dedication to my success, have profoundly impacted my personal and professional growth. I deeply appreciate how considerate and understanding you have been, making this journey not only productive but also reassuring and empowering.

I extend my heartfelt thanks to the surgeon gynecologist Dr. Fabien VIDAL, who have helped in gathering the data essential to this work and validating the different results.

I am also deeply grateful to the esteemed members of my PhD committee, Professors Su RUAN, Jean-Louis DILLENSEGER, Abderrahim HALIMI, Julian TACHELLA, Oumaima EL MANSOURI, Mireille GARREAU and Petar DJURIC. Thank you for taking the time to review my work, for your constructive comments, insightful questions and for making my defense a rewarding experience.

I would like to thank Corinne MAILHES for warmly welcoming me into TéSA Laboratory and providing me with an inspiring and supportive environment to conduct this PhD. Your kindness made this journey truly memorable. I also deeply appreciate you sharing some posters with me. Your support meant a lot during those moments.

I am also deeply thankful to all the wonderful people I had the pleasure of working with at TéSA. A special shout-out to my office mates, Valérian, Hamish and Jihanne, for sharing the ups and downs of PhD life with me. To Valérian, thank you for the many laughs, your basketball skills and our running joke about your level in Spanish. To Joan, Paul, Younes, Maurine, Jeff, Léa, Esteban, and Marta—my fellow PhD students, I wish you all the best of luck in your research and future endeavors.

I also want to thank Raoul, whose hilarious comments during lunch kept us all entertained. Kareth, you've been such a supportive presence in the lab, always checking in on me and sharing so many jokes that kept me going. It's been a real pleasure working alongside you. Samy, thank you for your genuine interest in my work and for always being so kind and approachable throughout this journey.

I would like to thank Victor, who has since left the lab, for always making sure I was included when he baked us pastries. I especially appreciated how you would make a mini gelatine-free cake just for me, so I could enjoy it too. Your thoughtfulness truly made a difference, and I'll always remember those kind gestures!

I started this journey with Hamish, Evelyne, and Gaston, and it's great to see us all reaching the finish line together. Isabelle, thank you for always greeting me with a warm smile—it truly brightened my days. Patrice, despite your tendency to forget names, you always made an effort to remember mine and showed genuine interest in my work, for which I'm grateful. Philippe, your helpfulness and kindness have been a constant source of encouragement. Lorenzo, thank you for your kindness and support—it has meant a great deal to me. And Wafae, even though you only come once a week, we quickly became good friends, and I'm glad we crossed paths.

To everyone in the lab, whether mentioned here or not, thank you for being part of this experience and making it all the more memorable. Your support and good spirit made this journey something I will always cherish.

I would now like to turn my appreciation to my personal circle. I would like to thank my friends and everyone who supported me along the way, but a special mention goes to those closest to me. Nisrine, thank you for always being up for a café or a shopping spree, and for saving me from buying all the random, unnecessary things I tend to pick up along the way. Abderrazak, thank you for always being there, whether it's for a last-minute airport run or for patiently answering all of my questions about life's little struggles. And Loutfi, thank you for being such a good sport about being the punchline of half of my jokes. You three are simply the best, and I couldn't have made it through this rollercoaster without your humor (or mine?), support, and friendship.

Finally, I would like to extend my deepest gratitude to my family, whose unwavering love and support have been a constant source of strength throughout this journey. I am incredibly grateful to my big family for always being so proud of me and for making me feel valued every step of the way. Words can't truly capture how thankful I am to my parents, who have always been there for me, offering their encouragement, understanding, and endless support—no matter the challenges. To my brothers, Otmane and Yahya, thank you for being a source of joy and for always having my back, reminding me that family is everything.

Résumé

L'endométriose est une maladie gynécologique chronique de la femme qui se caractérise par le développement d'un tissu semblable à la muqueuse utérine en dehors de l'utérus, colonisant d'autres organes avoisinants comme les ovaires, les trompes utérines ou dans des cas plus rares le côlon. Elle entraîne des douleurs abdominales et pelviennes, une fatigue chronique et un risque accru d'infertilité. Le diagnostic de l'endométriose repose sur deux modalités d'imagerie médicale, à savoir l'échographie (US) et l'imagerie par résonance magnétique (IRM). Selon le stade de maladie, la chirurgie laparoscopique s'avère être l'unique traitement efficace contre l'endométriose. Outre le diagnostic, les images US et IRM sont utilisées pour localiser précisément les lésions et leur profondeur d'infiltration avant la chirurgie. L'image US, effectuée par voie intravaginale pour cette application, fournit des détails internes fins des structures imagées grâce à sa haute résolution spatiale, mais a un champ de vision limité et un faible rapport signal/bruit. L'IRM, en revanche, offre un large champ de vision avec un bon rapport signal/bruit mais une résolution spatiale plus faible. Par conséquent, des repères anatomiques à l'échelle millimétrique seraient sous-évalués lors de l'utilisation de cette modalité seule. Construire une image rassemblant les avantages des deux modalités (bon contraste et bon rapport signal sur bruit) est d'un grand intérêt pour aider à la chirurgie de l'endométriose. Dans les applications pratiques, les examens US et IRM sont effectués séparément, ce qui donne des séries d'images US 2D et des volumes IRM 3D non-alignés.

L'objet de cette thèse de doctorat est d'abord de proposer un algorithme de recalage d'images IRM 3D et US 2D. Le but de ce recalage est d'extraire la coupe IRM qui ressemble le plus à l'image US, maximisant un critère de similarité adapté. Le recalage prend en compte une transformation rigide globale caractérisée par des paramètres de rotation et de translation qui est associée à une déformation locale basée sur des fonctions B-spline. Il permet une mise en correspondance plus précise entre les images, permettant d'exploiter également les déformations géométriques locales au sein de l'image.

Dans un deuxième temps, un modèle de fusion 2D/2D est proposé pour les images IRM et US. La méthode est basée sur un problème inverse, réalisant une super-résolution de l'image IRM et un débruitage de l'image US. La relation entre les niveaux de gris des deux images a été modélisée dans la littérature par une fonction polynomiale. Dans cette thèse, nous étudions l'intérêt potentiel de remplacer cette fonction polynomiale par une transformation non-paramétrique construite à

partir de la théorie des espaces de Hilbert à noyaux reproduisants. L'image fusionnée obtenue avec cette méthode rassemble les avantages des deux modalités, et présente un contraste plus net que lorsqu'on utilise un polynôme. Un autre avantage significatif en faveur de la transformation à base de noyaux est qu'elle n'est pas directement liée à la direction de propagation du scan US, qui n'est pas facile à obtenir dans les applications pratiques. Le prix à payer avec l'approche proposée est sa complexité. Le modèle peut nécessiter l'estimation de quelques centaines de milliers de paramètres en fonction de la taille de l'image et des patchs choisis.

Nous proposons un second modèle de fusion basé sur le filtrage guidé. L'image fusionnée est obtenue grâce à une pondération des images Base et Détail calculées à partir des images IRM et US. Les poids attribués à l'image US tiennent compte de la présence de bruit de speckle, tandis que les poids attribués à l'image IRM permettent d'améliorer le contraste de l'image fusionnée.

L'intérêt des modèles proposés est analysé au moyen de tests quantitatifs et qualitatifs effectués sur un ensemble de données varié, incluant des images synthétiques, des images d'un fantôme expérimental et des données réelles.

Abstract

Endometriosis is a chronic gynecological disease affecting women of childbearing age which is characterized by the development of tissue similar to the uterine lining (the endometrium) outside the uterus, colonizing other nearby organs such as the ovaries, the fallopian tubes or, in rarer cases, the colon. This tissue is influenced by hormonal changes during subsequent menstrual cycles, leading to abdominal and pelvic pain, chronic fatigue and an increased risk of infertility. The diagnosis of endometriosis is based on two medical imaging modalities, namely ultrasound (US) and magnetic resonance imaging (MRI). Depending on the stage of the disease, laparoscopic surgery proves to be the only effective treatment for endometriosis. Besides their use for diagnosis, US and MRI images are used to identify the precise location of lesions and their depth of infiltration before surgery. The US image, performed intravaginally for this application, is a high spatial resolution modality that provides fine internal details of the imaged structures. This modality has some limitations, including a limited field of view and a low signal-to-noise ratio. On the other hand, MRI offers a large field of view of the patient's anatomy with a good signal-to-noise ratio but with relatively low spatial resolution. Therefore, precise anatomical landmarks at the millimeter scale would be undervalued when using this modality alone. In this context, producing an image bringing together the advantages of both modalities (good contrast and good signal-to-noise ratio) would be of great interest. In practical applications, US and MRI examinations are performed separately, resulting in unaligned 2D US images and 3D MRI volumes.

The first aim of this PhD thesis is to propose a slice-to-volume registration algorithm of 3D MRI and 2D US images. The goal of this registration would be to extract the MRI slice that best resembles the US image, maximizing an adapted similarity criterion. The registration takes into account a global rigid transformation characterized by rotation and translation parameters which is associated with a local deformation based on B-spline functions. The latter will allow more precise matching between images, making it possible to exploit local geometric deformations within the image.

Secondly, a 2D/2D fusion model is proposed for MRI and US images. The method is problem-based inverse, achieving super-resolution of the MRI image and denoising of the US image. The relationship between the gray levels of the two images has been modeled in the literature by a polynomial function. We study the potential interest of replacing this polynomial function by a non-parametric transformation constructed from the theory of Hilbert spaces with reproducing kernels. The fused image obtained with this method combines the advantages of both

modalities, and presents a sharper contrast than when using a polynomial. Another significant advantage in favor of the kernel-based transformation is that it is not directly related to the propagation direction of the US scan, which is not easy to obtain in practical applications. The drawback of the proposed approach is its complexity. The model may require the estimation of a few hundred thousand parameters depending on the size of the image and the patches chosen.

We propose a second fusion model based on guided filtering, which consists of separating images into base and detail layers, calculating specific weights, and then fusing them. The fused image is obtained by weighting the Base and Detail images of the MRI and the US. The weights assigned to the US image take into account the presence of speckle noise, while the weights assigned to the MRI make it possible to improve the contrast of the fused image.

The interest of the proposed models is analyzed by means of quantitative and qualitative tests carried out on several datasets, including synthetic images, images of an experimental phantom and real data.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Contents

1.1 Endometriosis

Endometriosis is a chronic inflammatory medical condition that occurs when tissue similar to the lining of the uterus (endometrium) grows outside of it. These growths, also known as lesions or nodules, are most frequently found on the ovaries, fallopian tubes, the intestines, or the bladder. In rare cases, this endometrial-like tissue can even appear on other internal organs (Fig. 1.1). This aberrant tissue behaves like normal endometrial tissue, responding to the hormonal changes of the menstrual cycle by growing, thickening, and then breaking down and bleeding. However, unlike the tissue within the uterus, this displaced tissue has no way to exit the body. As a result, it becomes trapped, leading to inflammation, the formation of scar tissue, and adhesions. These adhesions can cause organs to stick together, creating a variety of complications [Eskenazi 1997].

Endometriosis affects approximately one in ten women of childbearing age (from the first menstrual period until menopause), i.e. about 1.5 million women in France, according to the Ministry of Health, and almost 200 million worldwide, according to the World Health Organization (WHO) [Fazleabas 2002].

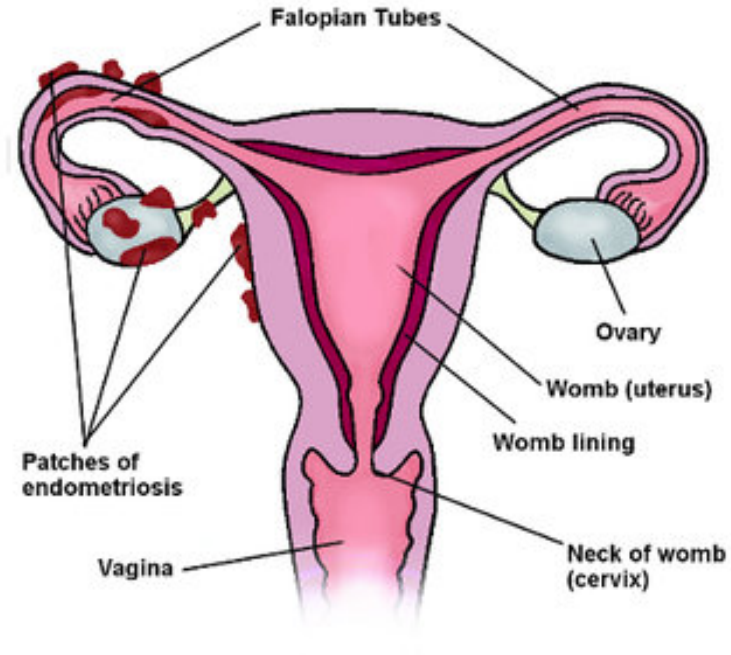


Figure 1.1: Anatomy of a uterus with endometriosis [Endometriosis-UK].

1.1.1 Symptoms of endometriosis

The symptoms of endometriosis vary widely among affected individuals, both in type and severity. The most common symptoms are pelvic pain and infertility. Pelvic pain is often chronic and can be particularly severe during menstruation, a condition known as dysmenorrhea [Harada 2013]. However, this pain is not limited to menstrual periods; it can also occur at other times, causing significant discomfort and disruption to daily life. Pain during or after sexual intercourse (Dyspareunia) is another frequent symptom [Denny 2007]. Additionally, women with endometriosis may experience menstrual irregularities such as heavy bleeding (menorrhagia) or bleeding between periods. Endometriosis can also affect bowel and bladder function, leading to painful bowel movements or urination, especially during menstruation. Other symptoms include fatigue, depression and anxiety, further impacting the quality of life for those with the condition [Ramin-Wright 2018].

Regarding infertility, one third of women with infertility problems have endometriosis and among women with endometriosis, approximately 40% are infertile [Bulletti 2010]. The mechanisms connecting endometriosis to infertility are still not fully understood. In fact, the inflammation caused by the endometrial-like lesions can directly damage both eggs and sperm, reducing their viability and function. Additionally, the formation of scar tissue and adhesions can block the fallopian tubes, preventing the egg and sperm from meeting, or impair the mobility of the fallopian tubes, affecting the transport of the fertilized egg to the uterus (Fig. 1.2).

Furthermore, endometriosis can disrupt the hormonal environment within the

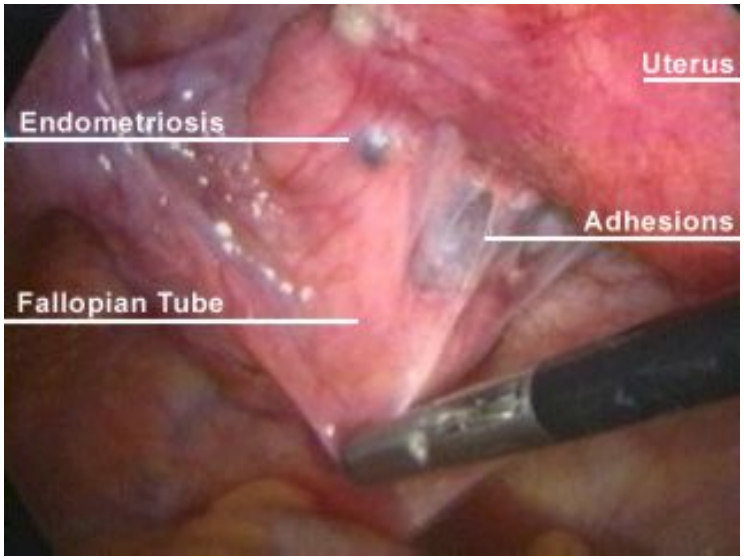


Figure 1.2: An example of endometriosis lesions and adhesions that attach the fallopian tube to the uterus [Fertility 2024].

pelvis. This alteration in hormone levels can negatively impact ovulation, the process by which an egg is released from the ovary. It can also interfere with the implantation of an embryo into the uterine lining, a crucial step for a successful pregnancy.

The average time to detect endometriosis is between seven and ten years during which patients suffer from physical and psychological consequences [Ghai 2020]. This huge delay in diagnosis can be explained by the variability of symptoms and their overlap with other conditions. Women who experience pain during menstruation are usually not immediately concerned, and unfortunately neither are those around them or medical personnel. The tendency to dismiss menstrual pain as a routine part of being a woman can lead to significant delays in seeking and receiving appropriate medical attention. Moreover, there are numerous other potential causes of period pain, painful intercourse, and bowel movements that need to be ruled out, which complicates the diagnostic process. Conditions such as irritable bowel syndrome, pelvic inflammatory disease, and ovarian cysts can present with similar symptoms, leading to misdiagnosis or underdiagnosis.

A lack of awareness about endometriosis among both medical professionals and the wider population contributes to the lengthy time to diagnosis. For the general population, there is often a limited understanding of endometriosis, which can lead to women normalizing their pain and not advocating for themselves as strongly as they might for more widely recognized conditions [Fazleabas 2002].

1.1.2 Causes of endometriosis

The exact cause of endometriosis is not fully known, but several theories exist. One theory, known as retrograde menstruation [Fallas 1956], suggests that menstrual blood flows backward through the fallopian tubes into the pelvic cavity instead of leaving the body, causing endometrial cells to implant and grow outside the uterus. Another theory involves embryonic cell transformation, where hormones such as estrogen may transform embryonic cells into endometrial-like cell implants during puberty. Additionally, surgical scars from procedures like hysterectomy or Caesarean-section can provide sites for endometrial cells to attach and grow. Immune system disorders may also play a role, as problems with the immune system can make the body unable to recognize and destroy endometrial-like tissue growing outside the uterus, allowing the condition to persist and worsen over time.

Several factors can increase the risk of developing endometriosis. Women who have never given birth are at higher risk, as are those who started menstruating at an early age or go through menopause at an older age. Short menstrual cycles (less than 27 days), heavy menstrual periods lasting longer than seven days, and higher levels of estrogen or greater lifetime exposure to estrogen also increase the risk [Darrow 1993]. A low body mass index (BMI) and a family history of endometriosis (having close relatives such as a mother, aunt, or sister with the condition) are additional risk factors [Zondervan 2001].

1.2 Endometriosis diagnosis and treatment

The diagnosis of endometriosis typically begins with a detailed medical history and a physical examination. A pelvic exam is performed to check for painful spots, irregular growths or cysts that can be palpated. However, small areas of endometriosis can't be felt during this exam. To enhance the accuracy of the diagnosis, medical imaging techniques are employed. The most commonly used imaging techniques are Transvaginal Ultrasound (TVUS) and Magnetic Resonance Imaging (MRI), both of which can provide detailed images of the pelvic region and help identify abnormalities [Smith 1991] [Zanardi 2003]. Computed tomography (CT) scanning is not typically performed in the radiologic evaluation of endometriosis because the appearance of endometriosis is nonspecific. If CT scanning is performed, endometriomas appear as cystic masses [Weinfeld 1998]. At this stage, doctors make a "clinical diagnosis" based on the medical history, physical exam findings and imaging results. Although this is not a definitive diagnosis, it allows the patient to begin prescribed medications right away without needing an invasive surgical diagnostic procedure, to avoid delays in treatment.
The only conclusive way to diagnose endometriosis is through laparoscopy. This surgical procedure allows the surgeon to directly visualize the endometrial tissue and, if necessary, take a biopsy for further examination.

The condition is then classified into four stages: Stage I (minimal), Stage II (mild), Stage III (moderate), and Stage IV (severe). The determination of these stages is based on several factors, including the location, size, and depth of organ infiltration by the endometrial tissue [Parazzini 1995].
There is no cure for endometriosis, but depending on the stage of the disease, several treatments can help manage symptoms:

- **Pain Medication**: Medications such as pain relievers, can help alleviate the chronic pain and discomfort associated with the condition. These medicine include nonsteroidal anti-inflammatory drugs and acetaminophen [Brown 2017].
- **Hormone Therapy**: Hormone treatments, including birth control pills, GnRH agonists, progestin therapy, and aromatase inhibitors, can help regulate or suppress the hormonal fluctuations that trigger the growth and shedding of endometrial-like tissue, thereby reducing the proliferation of endometriosis and alleviating symptoms [Tosti 2017]. Combined (Estrogen-progestin) contraceptives are often the first-line treatment [Vercellini 2016], they work by suppressing the ovaries, which may slow the progression of the endometriosis. Danazol helps reduce endometriosis-caused pain. It resembles testosterone and therefore can cause side effects, including acne, hirsutism (excessive hair growth in women), and a deepening of the voice. Because of these side effects, danazol is not widely used [Barbieri 1982].
- **Surgical Intervention**: Conservative surgery (Laparoscopy) removes endometriosis tissue. It aims to remove as much endometriosis as possible while preserving the uterus and ovaries [Duffy 2014].
- **Hysterectomy**: It consists in the ablation of the uterus [Martin 2006]. Hysterectomy is the last resort to ease pain for women with endometriosis who still experience symptoms despite other treatments and who do not intend to become pregnant in the future [Uccella 2016].

1.3 Laparoscopy

Laparoscopy, also known as minimally invasive surgery (MIS), is a diagnosis and treatment surgical procedure that allows surgeons to access the inside of the abdomen and pelvis without having to make large incisions in the skin. It is a commonly used technique in modern surgery due to its advantages in terms of reduced recovery time, minimal scarring, and decreased postoperative pain [Duffy 2014]. This surgery is administrated under general anesthesia, ensuring the patient remains unconscious and pain-free throughout. The patient is positioned on the operating table, generally lying on their back, and the abdominal area is cleaned and sterilized. The surgeon then makes a small incision, usually around the navel, through which a cannula (a thin tube) is inserted. This cannula is used to inflate the abdomen with carbon dioxide gas, creating a working space

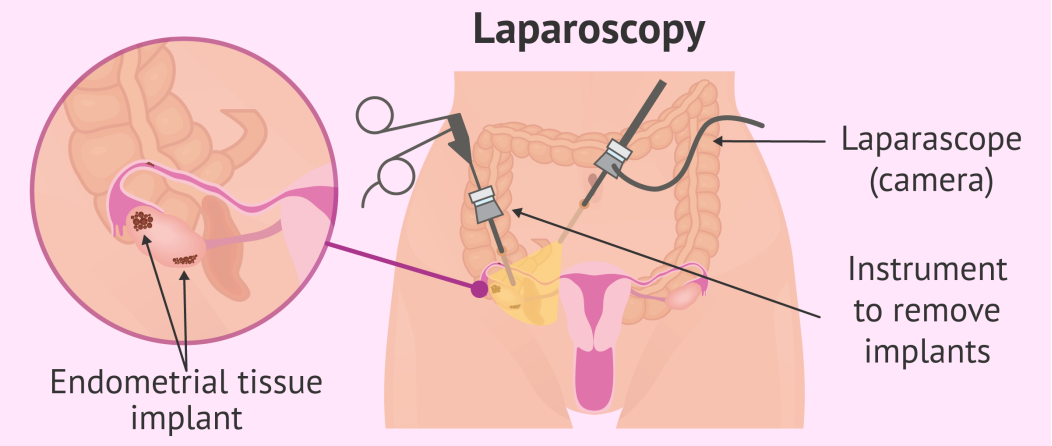


Figure 1.3: Laparoscopy protocole for endometriosis [Larrea 2024].

by lifting the abdominal wall away from the internal organs. A laparoscope, a long, thin tube equipped with a high-intensity light and a high-resolution camera at its end, is then inserted through this incision. The camera transmits images to a video monitor, providing the surgeon with a detailed real-time view of the pelvic cavity. One or more additional small incisions are made to insert specialized instruments to manipulate, cut or remove tissues as needed (Fig. 1.4). After completing the procedure, the instruments are withdrawn, the carbon dioxide gas is allowed to escape and the small incisions are finally closed with stitches (Fig. 1.3).

This endometriosis surgery is particularly sensitive and challenging due to several critical factors. One significant aspect is that many women undergoing this surgery are of childbearing age. The primary goal of the surgeon is therefore to meticulously remove as much of the endometrial lesions as possible to alleviate pain and other symptoms, while ensuring the uterus and ovaries remain intact and functional, to preserve the patient's fertility and ability to conceive in the future.

This balance between thoroughness and preservation is delicate, and it is often impossible to remove 100% of the lesions. Some may be microscopic, deeply infiltrated, or located in areas that are too risky to operate on without causing significant harm. This incomplete removal can lead to a recurrence of symptoms. Studies suggest that endometriosis can recur in up to 40-50% of women within five years after conservative surgery [Guo 2009].

The precision of a laparoscopy procedure relies heavily on the accuracy of the pre-operative diagnosis, which is primarily achieved through imaging techniques such as MRI and US. These imaging modalities are employed to detect and evaluate the lesions, their precise locations, and the depth of their infiltration into surrounding tissues before the surgery. Such detailed information is essential for effective surgical planning, and doctors rely on these images to plan the removal of the lesions [Craig 2020]. However, once the surgery begins, the surgeon operates

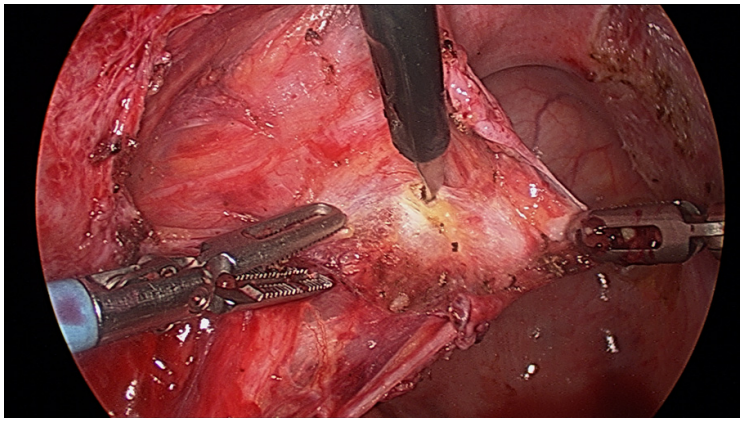


Figure 1.4: Laparoscopic excision of a moderate endometriosis lesion [Gynaescope 2015].

without direct access to these images with no possibility of checking them back if a nearby monitor is not available. The real-time visualization provided by the laparoscope during the procedure offers a different perspective, and the surgeon must rely on their memory and understanding of the pre-operative data. This is due to time constraints that make it impractical to leave the sterile field, access preoperative images and then return to the precodure. Surgeons already manage a high cognitive load during procedures, making numerous decisions in real-time. Having to mentally switch between real-time operative views and preoperative images is inefficient as it adds to this cognitive load. Also, prolonged anesthesia exposure can increase the risk of complications such as infections.

1.4 Imaging modalities

In this section, we will delve into the imaging modalities used for diagnosing endometriosis, specifically MRI and US. Having established their critical role in enhancing the precision of laparoscopy, we will first provide an overview of these imaging techniques and then focus on their application in the context of endometriosis.

1.4.1 US Imaging

Ultrasound imaging uses high-frequency sound waves to view inside the body. The frequencies used are higher than those audible to the human ear, typically ranging between 2 to 20 Megahertz. These high-frequency sound waves are generated by a transducer or probe, which contains special ceramic crystal materials called piezoelectrics. When an electrical current is applied to these crystals, they vibrate and produce ultrasound waves. When the returning echoes strike the crystals, they generate an electrical signal that is then converted into an image [Genovese 2016]. The transducer emits these sound waves into the body, where they travel through tissues. To ensure efficient transmission, a gel is applied between the transducer

and the skin to eliminate air gaps, which would otherwise cause significant loss of sound energy. The ultrasound waves pass through the body and reflect off different internal structures before returning to the transducer as echoes.

These echoes vary in intensity depending on the density and type of tissues they have encountered. Strong echoes return from dense structures like bone, while weaker echoes return from softer tissues. The transducer receives these echoes and sends them to the ultrasound scanner, which processes the data to create an image (Fig. 1.5). Using the the speed of sound and the time it took the echoes to return, the machine constructs a two-dimensional image of the internal structures, where the brightness (or echogenicity) of each pixel represents the intensity of the echo. Dark areas typically indicate fluid-filled structures (like cysts or soft tissues), while brighter areas indicate denser structures or greater acoustic impedance, which reflects more ultrasound waves back to the transducer (like bone or tumors).

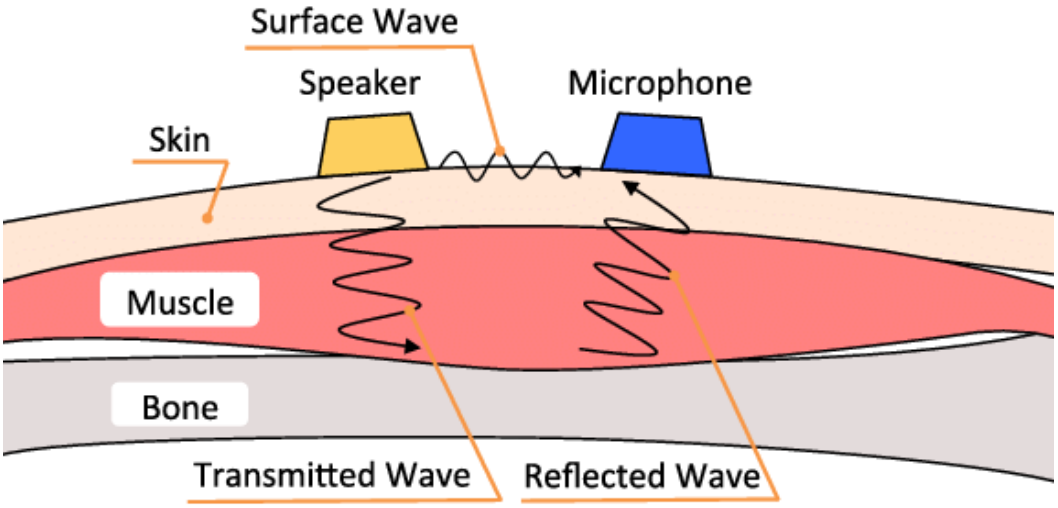


Figure 1.5: Mechanism of US Imaging [Watanabe 2017].

Ultrasound imaging is used for many medical procedures, some examples include nephrology [Singla 2022] (to detect kidney stones for instance), gynecology, gastroenterology and cardiology [M Thijssen 2014].

1.4.1.1 Interaction of Ultrasound Waves with Tissues

The US image is formed by transmitting pulses of high-frequency sound waves into the body and detecting the reflected echoes. As these sound waves travel through different tissues, they encounter boundaries between tissues of varying densities. Many phenomena can occur during their propagation that affect the quality of the resulting image [Chan 2011].

- **Reflection**: It occurs when ultrasound waves encounter a boundary between tissues with different acoustic impedances. At these boundaries, part of the

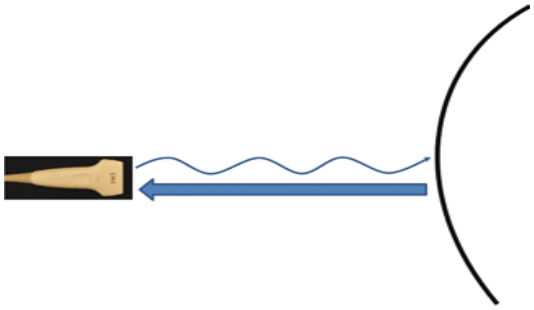


Figure 1.6: Illustration of wave reflection [Ultrasonography 2016].

ultrasound wave is reflected back towards the transducer while the rest continues to travel deeper into the body (Fig. 1.6). The intensity of the reflected wave depends on the difference in density between the two tissues. For instance, the boundary between muscle and bone produces a stronger echo compared to the boundary between muscle and fat. These returning echoes are detected by the transducer and converted into electrical signals, which are then processed to form the ultrasound image.

The changing intensities of these echoes create the different shades of gray seen on the image. Reflection is the primary mechanism through which ultrasound images are formed, providing detailed information about the internal architecture of the body.

- **Refraction**: It happens when ultrasound waves pass through tissues of varying densities, causing the waves to bend at the interface due to changes in the speed of sound (Fig. 1.7). This bending happens because the speed of sound changes as it moves from one type of tissue to another with different acoustic properties. Refraction can lead to displacement or distortion of the ultrasound waves, resulting in artifacts in the resulting image.

- **Scattering**: Is the diffusion of the US beam in many directions. It happens when the ultrasound waves encounter irregular structures within tissues, by far smaller than the incident wavelength, causing them to disperse in multiple directions rather than reflecting directly back to the transducer (Fig. 1.8). These structures, known as scatterers, include red blood cells or the elementary cells within tissues or organs. The scattered waves can interact with each other, leading to constructive and destructive interferences. Constructive interference occurs when the scattered waves coincide in phase, amplifying the echo signal, while destructive interference happens when the waves are out of phase, reducing the echo signal. This scattered waves result in the typical speckle noise of US images.
  Speckle noise gives a grainy texture to the resulting image, that can obscure fine details and compromise the clarity of the visual information.

- **Attenuation**: It refers to the gradual loss of ultrasound wave energy as it

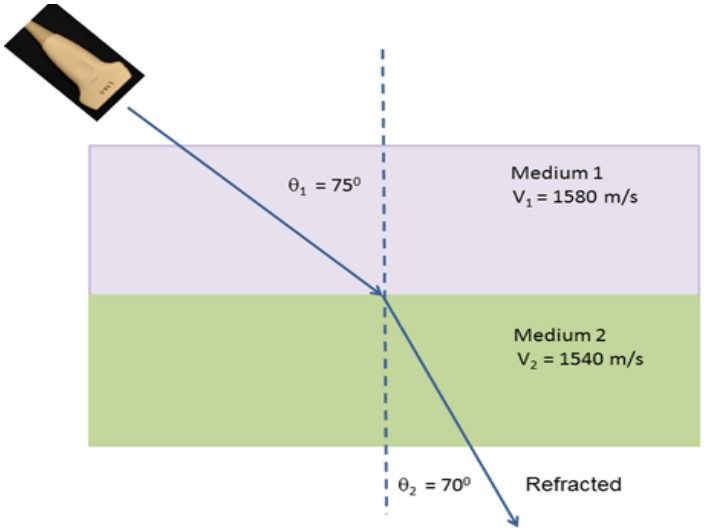


Figure 1.7: Illustration of wave refraction [Ultrasonography 2016].

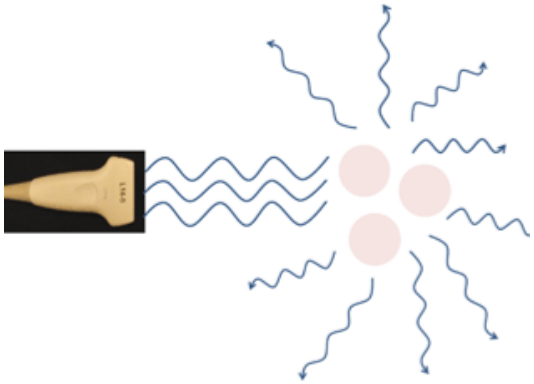


Figure 1.8: Illustration of wave scattering [Ultrasonography 2016].

travels through tissues. It is influenced primarily by two factors: the frequency of the ultrasound waves and the properties of the medium through which they travel. Higher frequency waves experience greater attenuation because they are absorbed and scattered more readily by tissues compared to lower frequency waves. Consequently, for imaging deeper structures, lower frequency transducers (typically in the range of 3-5 MHz) are utilized. These lower frequencies penetrate tissues more effectively, reducing the impact of attenuation, but they come at the expense of spatial resolution.

1.4.1.2 Ultrasound Modes

Different techniques are used to capture and display information in US imaging, referred to as modes:

- **A-mode**: A-mode (Amplitude mode) is one of the simplest ultrasound imag-

ing techniques. It displays the amplitude of the returning echoes as a function of time or depth along a single line. When displayed, on the X axis we see depth, whilst on the Y axis we see the amplitude of the echo response (Fig. 1.9). When the waves are reflected from different densities of materials, we see distinct peaks and troughs, allowing us to assess how deep each layer of material is. A-mode is limited as it only provides a one-dimensional view and does not show the spatial relationships about structures.

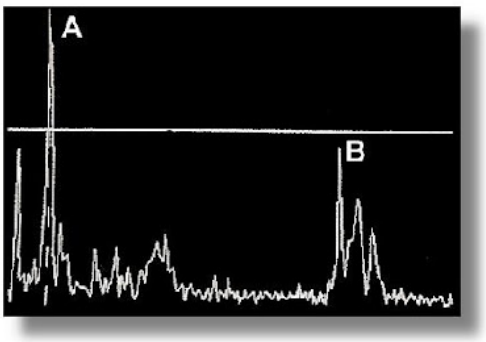


Figure 1.9: A-mode.

- **B-mode**: B-mode (Brightness mode) is a two-dimensional image (Fig. 1.10).

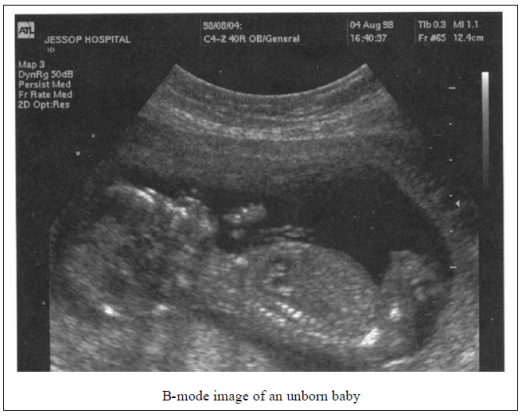


Figure 1.10: B-mode.

It represents the envelopes of the received radio-frequency (RF) signals following a logarithmic compression operation, defined by the equation:

$$y_{\mathrm{US}} = b \log(|r_{\mathrm{IQ}}|) + g \tag{1.1}$$

where b and g are constants for linear gain, and y_{US} denotes the B-mode US image. r_{IQ} is the phase and quadrature signal obtained from the raw signal

r_{RF} captured by the transducer. For the remainder of this work, the focus will be on the B-mode US image, which will be referred to simply as the US image.

- **M-mode**: M-mode (Motion mode) allows us to view the motion of structures over time by recording a single line of ultrasound data as a function of time (Fig. 1.11). It is primarily used in cardiology to assess the motion of heart valves and the movement of cardiac walls.

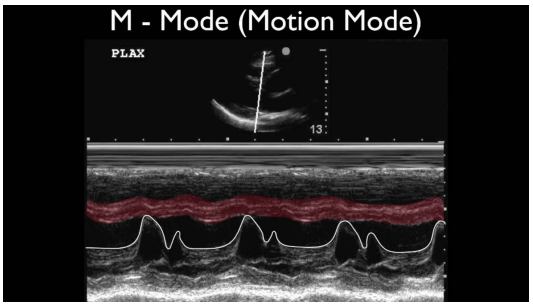


Figure 1.11: M-mode.

- **Doppler mode**: It assesses blood flow and the movement of fluids within the body by measuring changes in the frequency of the reflected ultrasound waves (Fig. 1.12).

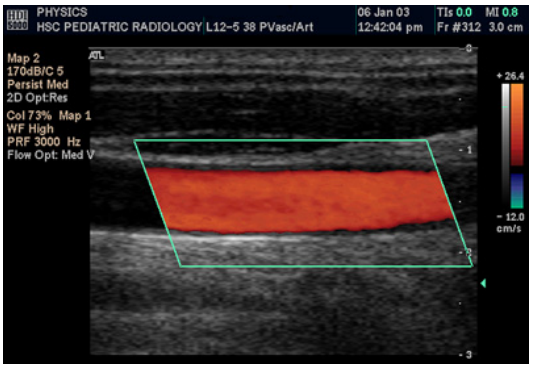


Figure 1.12: Doppler mode.

For more details about the US image formation, the reader is referred to the book [Sanches 2012].

1.4.2 MR Imaging

Magnetic Resonance Imaging is a sophisticated medical imaging technique used to visualize detailed internal structures and tissues within the body. MRI utilizes

strong magnetic fields and radiofrequency pulses to generate images of organs, tissues, and other internal structures without using ionizing radiation, unlike X-rays or CT scans (Fig. 1.13).

Figure 1.13: MR machine

MRI systems use powerful magnets, typically superconducting magnets, to generate a static magnetic field with strengths typically ranging from 1.5 to 3.0 Tesla, and sometimes even higher. The underlying principle of MRI is based on the phenomenon of nuclear magnetic resonance (NMR) [Young 1984] [Kinley 2015]. It involves the interaction between hydrogen nuclei and the magnetic field. When the strong magnetic field is applied, the hydrogen nuclei, which have a magnetic moment due to their spin, align with the field. This alignment creates a net magnetization in the direction of the magnetic field. The application of Radiofrequency (RF) pulses causes these nuclei to absorb energy and move to a higher energy state. Once the RF pulse is removed, the hydrogen nuclei return to their original alignment with the magnetic field, a process known as relaxation [Tang 2018]. During relaxation, the nuclei release the absorbed energy in the form of radiofrequency signals. These emitted signals vary based on the tissue types and their chemical environments, allowing the MRI system to capture detailed images of different tissues and structures within the body. This process of alignment and signal production is central to the generation of high-resolution MRI images. Fig. 1.14 explains the

mechanism of magnetic resonance imaging.

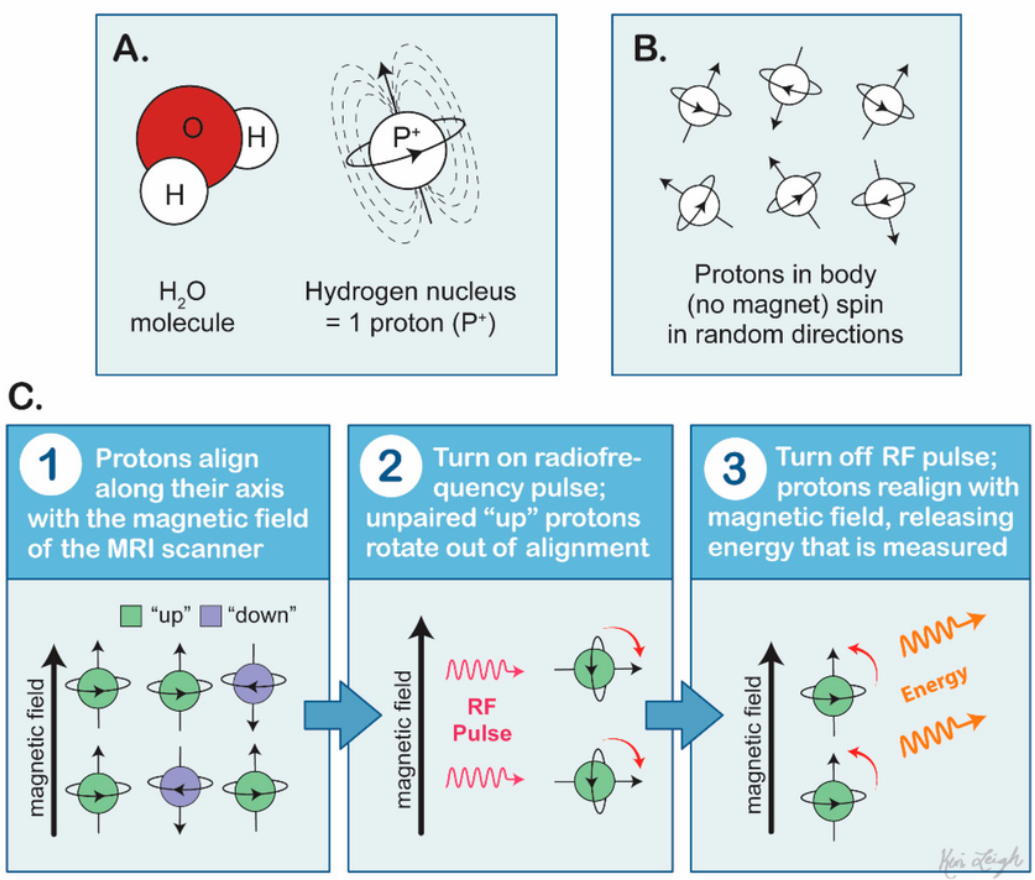


Figure 1.14: Mechanism of MR Imaging [Fordham 2021].

RF pulses play a pivotal role in MRI by exciting the hydrogen nuclei within the body's tissues. Once the hydrogen nuclei are aligned with the strong magnetic field, RF pulses, which are short bursts of electromagnetic energy, are applied. These pulses are tuned to a specific frequency that matches the resonance frequency of the hydrogen nuclei. When the RF pulse is applied, it temporarily disrupts the alignment of the hydrogen nuclei, causing them to move out of alignment with the magnetic field. This excitation process is essential for generating the signals needed for imaging. After the RF pulse is turned off, the hydrogen nuclei begin to realign with the magnetic field, releasing energy in the form of radiofrequency signals. These signals are then detected by the MRI system and used to construct images of the internal structures.

MRI scanners are especially effective for imaging the body's soft tissues, as opposed to bony structures. They provide superior visualization of the brain, spinal cord, nerves, muscles, ligaments, and tendons compared to traditional X-rays and CT scans. This enhanced clarity makes MRI particularly valuable for assessing injuries and conditions involving soft tissues, such as those in the knee and shoulder.

As a result, MRI is frequently used to diagnose and evaluate injuries in these areas.

1.4.2.1 MRI Imaging sequences

Tissues can be characterized by two distinct relaxation times: T1 and T2. T1, known as the longitudinal relaxation time, represents the time constant that determines how quickly excited protons return to their equilibrium state. It measures the duration required for protons to realign with the external magnetic field after being disturbed. T2, or transverse relaxation time, describes the time constant that determines how rapidly excited protons lose phase coherence with one another. It indicates the time needed for spinning protons to become out of phase among the nuclei spinning perpendicular to the main field (Fig. 1.15):

- **T1-Weighted Images**: T1-weighted MRI images are characterized by their ability to provide high-resolution images where tissues with short T1 relaxation times (such as fat) appear bright, and those with longer T1 times (such as water) appear darker. This imaging mode is particularly useful for evaluating anatomical detail and differentiating between different types of tissues. T1-weighted images are often employed in neuroimaging to assess brain structure, identify lesions, and visualize anatomical features such as the gray and white matter. They are also useful in musculoskeletal imaging for evaluating the integrity of soft tissues and detecting changes in fat content.

- **T2-Weighted Images**: T2-weighted MRI images highlight tissues with long T2 relaxation times, making them appear bright, while tissues with shorter T2 times appear darker. This imaging mode is valuable for detecting fluid and edema because water-based tissues, like cerebrospinal fluid and inflammatory fluid, are prominently displayed. T2-weighted images are commonly used in diagnosing and monitoring conditions such as tumors, infections, and degenerative diseases, particularly in the brain and spinal cord. They are also useful in musculoskeletal imaging for visualizing joint effusions and soft tissue injuries.

- **Proton Density Images**: Proton density (PD) images are characterized by the contrast provided by the concentration of hydrogen nuclei in tissues. Unlike T1 and T2-weighted images, which are influenced by relaxation times, PD images emphasize differences in the density of hydrogen nuclei, providing a clearer distinction between tissues with varying hydrogen concentrations. PD imaging is particularly useful in musculoskeletal imaging for evaluating cartilage, ligaments, and tendons, as it provides clear anatomical detail without the contrast effects of T1 or T2 weighting.

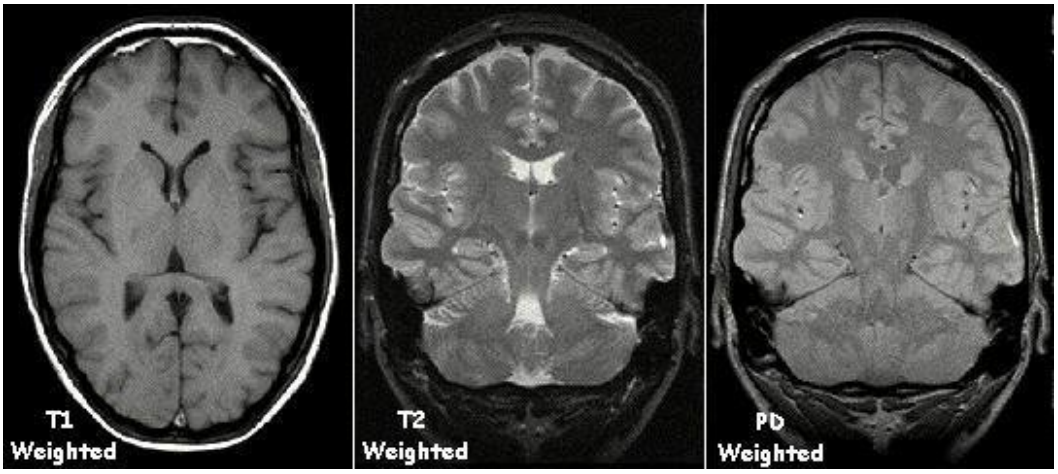


Figure 1.15: Example of T1-weighted, T2-weighted and PD-weighted MR images [Al-Majeed 2020].

Besides conventional MRI previously presented, other types of MRI techniques exist, each designed to provide specific information based on the area being examined and the condition being diagnosed. Functional MRI (fMRI) is a specialized imaging technique used to map brain activity by detecting changes in blood flow associated with neural activity. fMRI measures variations in the blood oxygenation level-dependent signal, which reflects changes in oxygenated and deoxygenated blood levels in active brain regions. It is commonly used in brain research to understand brain function, mapping areas responsible for specific cognitive tasks, and investigating neurological and psychiatric disorders [Gore 2003]. Another technique is Diffusion MRI (dMRI) that assesses the diffusion of water molecules within tissues, providing insights into tissue microstructure and integrity. Diffusion-weighted imaging (DWI), a subtype of dMRI, measures the direction and extent of water molecule movement, which can reveal abnormalities in the brain's white matter tracts and other tissues [Kiselev 2017].

1.4.3 US and MRI for endometriosis

Having established a foundational understanding of MRI and US, let us now explore their application in the diagnosis and treatment of endometriosis.
Most diagnostic US probes are typically placed on the skin for general imaging. However, to enhance image quality, especially in the context of pelvic conditions like endometriosis, probes may be inserted into the body via the gastrointestinal tract, vagina, or blood vessels. For endometriosis, transvaginal ultrasound (TVUS) is the most commonly used approach. This method allows for a closer and more detailed view of the pelvic organs. These organs include the uterus, the cervix, fallopian tubes and ovaries (Fig. 1.16).
TVUS plays a crucial role in diagnosing endometriomas (endometriosis cysts) and assessing deep endometriosis prior to surgery. It is particularly valuable for identi-

fying the extent of disease in women with a strong clinical suspicion of endometriosis. This imaging method is cost-effective, readily available, and requires no special preparation [Guerriero 2016]. The sonographer can detect deep infiltrating endometriosis, including adhesions and implants, and assess their size and location. Despite its advantages, TVUS has limitations such as significant speckle noise, which reduces the signal-to-noise ratio, and a restricted field of view with lower contrast. Enhancing the detection of deep infiltrating endometriosis through improved sonographic techniques could decrease the need for diagnostic laparoscopy, better inform management strategies, and ultimately improve patient quality of life.

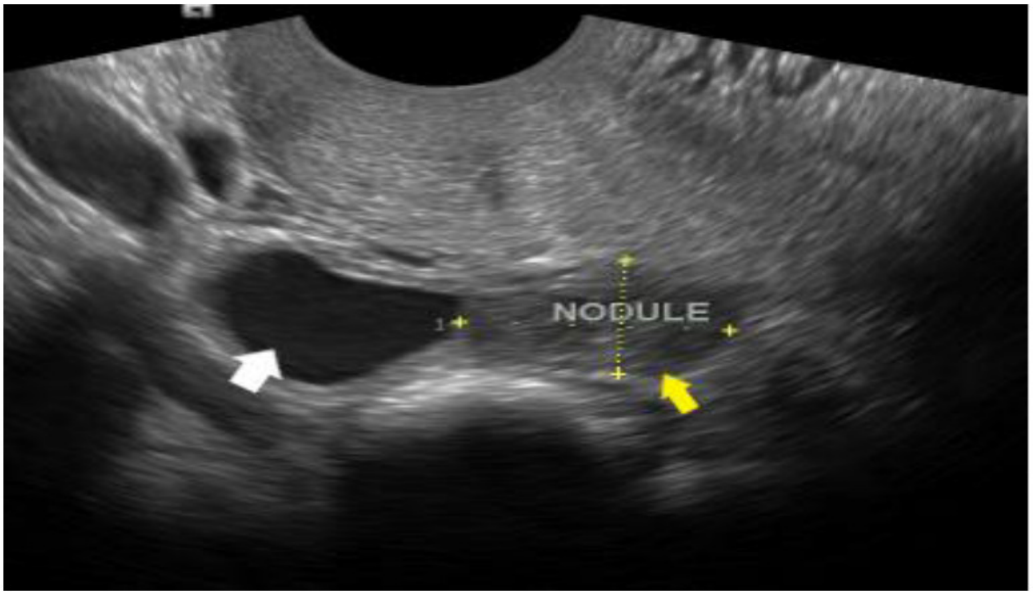


Figure 1.16: Example of a 2D TVUS image, showing a deep infiltrating endometriosis [Daniilidis 2022].

On the other hand, MR imaging produces three-dimensional detailed anatomical images. It can identify tissues through various parameter sequences and, unlike transvaginal ultrasound (TVUS), is not dependent on the operator's skill. Its wide field of view allows for comprehensive visualization of the entire pelvis, facilitating the evaluation and detection of lesions. MR images offer high contrast, enabling the creation of a precise map of endometrial implants (Fig. 1.17). However, MRI is not widely used for endometriosis diagnosis due to its high cost, limited availability, and lower spatial resolution, especially at the millimetric scale. MR images can provide sub-centimeter resolution. However, for early diagnosis, lesions can be very tiny and require millimetric scale resolution, making them subtle or undetectable.

Table 1.1 details the comparaison between MR and TVUS images.

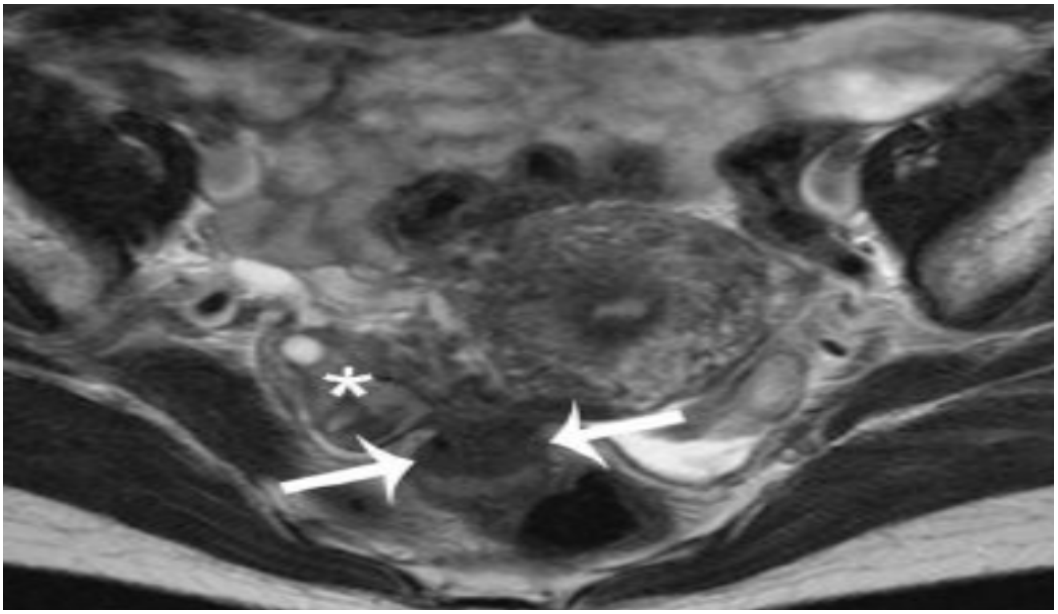


Figure 1.17: Axial MR T2 image, showing an endometriotic lesion infiltrating the rectum (arrows), creating adhesions with the right ovary (*) [Scardapane 2014].

Table 1.1: Comparison between MRI and Ultrasound (US)

Criteria	**MRI**	**US**
Expense	High	Low
Cost	Expensive	Affordable
Contrast	High	Low
Spatial Resolution	Limited	High
Field of View	Broad	Limited
Availability	Limited	Widely Available
Scan Time	Long	Short

1.5 Motivation and objectives

As we have explained in this detailed introduction, endometriosis is a complex and highly sensitive disease that affects millions of women worldwide. It can cause severe pain, infertility, and other debilitating symptoms. Accurate diagnosis and effective treatment planning are crucial for improving patient outcomes. Besides being used for diagnosis, US and MRI are essential tools for identifying endometrial implants and assessing their depth of infiltration in organs before surgery. These imaging techniques play a vital role in preoperative planning and determining the extent of the disease.

US imaging is a modality with high spatial resolution that provides fine internal details of structures. This makes it particularly effective for visualizing small anatomical features and detailed tissue characterization, which is critical in detecting and evaluating endometrial lesions. However, this modality is constrained by a limited field of view and a low signal-to-noise ratio, which can reduce the clarity and comprehensiveness of the images obtained.

In contrast, MRI offers a broad field of view of patient anatomy, providing a comprehensive overview of the pelvic region and beyond. It has a good signal-to-noise ratio, which enhances image clarity and contrast, making it useful for identifying deep and extensive endometrial implants. However, MRI's relatively low spatial resolution means that precise anatomical landmarks at the millimetric scale may be under-evaluated when using this modality alone. This limitation can be problematic when small lesions need to be identified and accurately mapped.

Fusing these two modalities appears particularly promising. Indeed, assembling all the critical information from both US and MRI modalities into a single informative image is expected to significantly improve pre-operative disease mapping and subsequent surgical approaches. The primary goal of this work is to generate an image that combines the high contrast of MRI with the fine resolution of US, thereby leveraging the strengths of both imaging techniques.

Previous work on the fusion of MRI and US images for endometriosis has primarily focused on the 2D/2D case [El Mansouri 2020]. Given that MRI provides a 3D volume, this PhD thesis goes a step further by first considering a 3D/2D registration algorithm, followed by the 2D/2D fusion process.

The first objective is to propose a registration model for 3D MR and 2D US images that will extract from the MR volume the image that best matches the US image. This involves aligning the MRI and US images to ensure that the anatomical features correspond accurately between the two modalities. Two fusion techniques are then proposed in the second part of this PhD thesis to generate a single informative image. The first technique involves a statistical model based on a least squares problem. The second technique employs guided filtering, which separates the images into base and detail layers, computes specific weights, and then fuses them.

This PhD thesis was conducted in collaboration with a gynecologist surgeon Dr. Fabien VIDAL from Clinique La Croix du Sud - Ramsay Santé, Toulouse,

France. He provided valuable clinical data essential for this research (experimental phantom, in vivo data), and played an integral role in validating the results.

1.6 Organization of the manuscript

The remaining of this thesis is organized as follows:

Chapter 2 provides a comprehensive overview of image registration, a fundamental technique in image processing and computer vision. It begins with a literature review that explores several methods proposed to address this challenging problem, with a particular focus on medical image registration. Following this, the main contribution of the chapter is presented, that is a general framework for registering 3D magnetic resonance (MR) and 2D ultrasound (US) images [Bennioui 2023] [El Bennioui 2023b]. This proposed registration accounts for a global rigid transformation, characterized by rotation and translation parameters, along with a local deformation based on B-spline functions. To validate the registration results, we apply a 2D-2D US-MRI fusion algorithm, generating a final image that combines the main characteristics of both MR and US images. The rationale behind this validation step is that well-aligned MR and US images will produce an accurate and high-quality fusion, while poorly matched images may result in artifacts or distortions.

The accuracy of the registration method is analyzed through quantitative and qualitative tests conducted on experimental phantoms, synthetic data generated from an in vivo MRI volume, and real data, with particular attention to endometriosis treatment.

Chapter 3 first presents state-of-the-art models and methods used in image fusion. It highlights the advantages, disadvantages, and challenges encountered in various applications. Next, the proposed fusion method for 2D Ultrasound (US) and 2D Magnetic Resonance (MR) images is presented. The goal is to create a hybrid image, referred to as the MARIUS image (MAgnetic Resonance Imaging & UltraSound), which combines the advantages of both modalities in the context of endometriosis diagnosis [El Bennioui 2023a].

This method is built upon the idea developed in [El Mansouri 2020], focusing on constructing two observation models associated with US and MR images to exploit the complementarity of both images in terms of resolution and contrast. For the MR image, a linear model formed by blurring and downsampling operators is proposed, motivated by its good performance in super-resolution. A despeckling model is employed for the US image to mitigate the effect of noise. Since US and MR modalities differ in their nature and content, there is no simple correspondence between the gray levels of these images. Therefore, we propose a new model using a non-parametric transformation defined using the theory of reproducing kernels and evaluating its benefits compared to the polynomial transformation investigated in [El Mansouri 2020].

The proposed fusion method is evaluated on an experimental phantom and synthetic data generated from an in vivo MRI volume, with specific attention to endometriosis treatment.

Chapter 4 explores a novel fusion method developed for combining magnetic resonance (MR) and ultrasound (US) images. The proposed method is based on guided filtering, leveraging the advantages of this technique to enhance the quality of fused images. The fusion process involves creating a weighted average of base and detail images derived from both MR and US images. Specifically, the weights for the US image address the challenge of speckle noise, while the weights for the MR image enhance the contrast of the fused result. The effectiveness of this method is assessed using both synthetic data, phantom data and real data demonstrating promising outcomes. The fused images produced by this method have the potential to improve visualization and support decision-making in endometriosis surgery, representing a significant advancement in the field of medical image fusion [El Bennioui 2024].

Chapter 5 concludes the present thesis and provides a guideline for future work.

List of Publications

1. Y. El Bennioui, O. El Mansouri, F. Vidal, A. Basarab and J.-Y. Tourneret, **Recalage et fusion d'images échographiques et par résonance magnétique au service de la chirurgie de l'endométriose**, in *Proc. Colloque Francophone de Traitement du Signal et des Images (GRETSI)*, Grenoble, France, Aug. 2023.

2. Y. El Bennioui, F. Vidal, A. Basarab and J.-Y. Tourneret, **Joint Registration and Fusion of 3D Magnetic Resonance and 2D Ultrasound Images for Endometriosis Surgery**, in *Proc. European Signal Processing Conference (EUSIPCO)*, Helsinki, Finland, Sept. 2023.

3. Y. El Bennioui, A. Bruguier, F. Vidal, A. Basarab and J.-Y. Tourneret, **Fusion of Ultrasound and Magnetic Resonance Images for Endometriosis Diagnosis: A Non-Parametric Approach**, in *Proc. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Los Suenos, Costa Rica, Dec. 2023.

4. Y. El Bennioui, A. Halimi, A. Basarab and J.-Y. Tourneret, **Fusion of Magnetic Resonance and Ultrasound Images using Guided Filtering: Application to endometriosis surgery**, in *Proc. European Signal Processing Conference (EUSIPCO)*, Lyon, France, Aug. 2024.

5. Y. El Bennioui, F. Vidal, A. Basarab and J.-Y. Tourneret, **A general framework for registration and fusion of Ultrasound and Magnetic Resonance Images for endometriosis laparoscopy**, International Journal of Computer Assisted Radiology and Surgery - (In preparation)

Chapter 2

2D/3D MR-US Image registration

Contents

This chapter aims to provide a comprehensive overview of image registration, a fundamental technique in the field of image processing and computer vision. Image registration involves aligning two or more images of the same scene or object to facilitate comparison, analysis, or integration of the information they contain. This process has applications across various domains, from remote sensing and astronomy to medical imaging. We start with a comprehensive literature review, that presents several methods that have been proposed to deal with this challenging problem, with a particular focus on medical image registration. Then, a general framework for the registration of 3D magnetic resonance (MR) and 2D ultrasound (US) images is presented. The proposed registration takes into account a global rigid transformation characterized by rotation and translation parameters, associated with a local deformation based on B-splines functions. A 2D-2D US-MRI fusion algorithm is

also applied to validate the registration results, generating a final image containing the main characteristics of the two MR and US images. The rationale behind this validation step is that if the MR and US images are well-aligned, the resulting fusion will be accurate and of high quality. However, if the images do not match well, the fusion process may introduce artifacts or distortions. The accuracy of the registration method is analyzed by means of quantitative and qualitative tests conducted on experimental phantom, synthetic data generated from an in vivo MRI volume, and real data with a specific attention to endometriosis treatment.

2.1 Literature review

Image registration involves aligning and merging data from multiple image sources into a unified coordinate system. It compares a source image with a reference image in order to identify the optimal geometric transformation that captures accurate spatial correspondence between them by optimising a registration criterion. Numerous applications requiring image registration have emerged, driving the community to develop increasingly precise and efficient strategies. In remote sensing, image registration is used to align and integrate images from different sensors and platforms, enabling the creation of multi-sensor/multi-temporal image datasets for land cover classification, change detection, and environmental monitoring [Bentoutou 2005]. It is also a key component in various computer vision tasks such as object recognition, scene understanding, and 3D reconstruction, where aligning images accurately is essential for feature matching and scene interpretation [Movahed 2020]. In the medical field, image registration techniques aid in the comparison and fusion of images from different modalities or acquisition times, assisting clinicians in diagnosing diseases, tracking disease progression, and evaluating treatment efficacy [Anila Satheesh 2021]. Moreover, the integration of preoperative and intraoperative images, provides surgeons with real-time guidance during surgical procedures, enhancing accuracy, and improving patient outcomes. In cardiac imaging, these techniques are employed for example to align cardiac images acquired from different modalities or cardiac phases, which helps in capturing and quantifying the motion of the cardiac cycle of the heart and myocardial viability [Atehortúa 2019]. Image registration is also extensively used in neuroimaging to facilitate the analysis of brain structures and functions by aligning brain images from different sessions [Rivaz 2015].

The primary goal of image registration techniques is to spatially or temporally align a moving image with a fixed image target. These registration algorithms typically consists of three main components:

- Transformation model: Rigid or elastic (non-rigid) to be performed on the moving image enabling the images to be registered with local geometric differences.

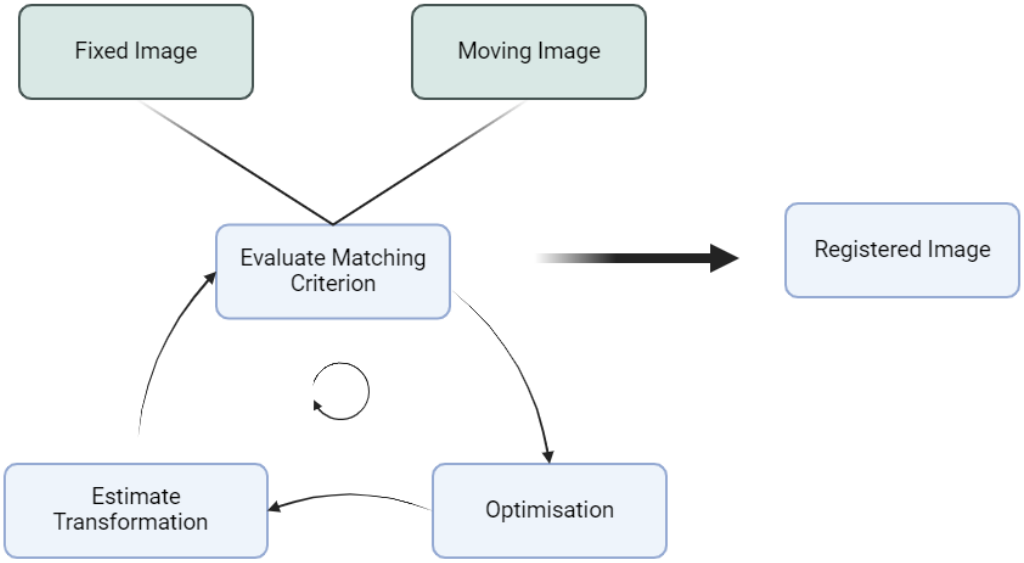


Figure 2.1: General image registration framework

- Cost function: Often referred to as a matching criterion or a (dis)similarity measure that quantifies the degree of alignment between the images.

- Optimization method: Estimates the optimal transformation to best minimize/maximize the defined matching criterion.

Figure 2.1 summarizes the iterative registration framework. It is typically performed by estimating a transformation using the model, applying it to the moving image and evaluating the cost function between the fixed and moving images. The cost function guides the algorithm in how to estimate a more accurate transformation for the next iteration. This process continues and optimizes until either the moving and fixed images are deemed aligned (reaching a local minimum in the cost function) or a predefined maximum iteration count is reached.

2.1.1 Matching Criterion

The matching criterion quantifies the alignment between the images. Registration methods can be classified into two categories, namely, intensity-based (iconic) methods, and feature-based (geometric) methods.

2.1.1.1 Intensity-based methods

Intensity-based approaches rely on the intensity information associated to pixels/voxels in the input images. They either consider the intensities themselves, or evaluate a statistical information from them, such as gradients, gray levels, etc. In

monomodal cases, i.e. the images are captured using the same modality, pixel/voxel intensity values corresponding to the same anatomical structures are highly correlated or even identical in both images. Therefore, the task of measuring similarity between the images is straightforward and simple measures can be used. Examples of these metrics include the sum of squared differences (SSD) [Fogtmann 2014], [Rueckert 2010], [Heldmann 2009]. It is defined as follows (for two vectorized images $\boldsymbol{X}$ and $\boldsymbol{Y}$):

$$\mathrm{SSD}(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{i \in \Omega} [X(i) - Y(i)]^2, \tag{2.1}$$

where Ω is the region of interest. Similar to SSD, sum of absolute differences (SAD) [Ferrante 2013] provides a measure of the overall difference between images, with smaller SAD values indicating greater similarity. These metrics operate under the assumption of a direct correspondence between the intensity values in both images.

Another metric used is the Pearson Correlation Coefficient (CC). It is a measure of the strength and direction of the linear relationship between two variables [Sarvaiya 2009], [Van den Elsen 1995]. Given a pair of images $\boldsymbol{X}$ and $\boldsymbol{Y}$, the Pearson correlation coefficient is defined as:

$$\mathrm{CC}(\boldsymbol{X}, \boldsymbol{Y}) = \frac{\mathrm{cov}(\boldsymbol{X}, \boldsymbol{Y})}{\sigma_{\boldsymbol{X}} \sigma_{\boldsymbol{Y}}}, \tag{2.2}$$

where $\mathrm{cov}(.,.)$ and σ are the covariance and standard deviation operators.

In the multimodal scenario, such simple metrics can present challenges due to the lack of obvious relationships between pixel intensities from different modalities. More sophisticated metrics have therefore been developed to leverage the statistical properties of the observed intensity values in images. The most popular is the mutual information (MI). Using mutual information for image registration was originally presented in [Wells 1996], and since then, it has become one of the most widely used registration cost functions [Sengupta 2022]. Its success largely comes from its probabilistic nature, which gives it robustness to noise and shifts in intensity. MI quantifies the "amount of information" provided by one random variable given that the other random variable is observed. It is defined as:

$$\mathrm{MI}(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{x \in \Omega_{\boldsymbol{X}}} \sum_{y \in \Omega_{\boldsymbol{Y}}} p(x, y) \log \left[\frac{p(x, y)}{p(x) p(y)} \right], \tag{2.3}$$

where $\Omega_{\boldsymbol{X}}$ and $\Omega_{\boldsymbol{Y}}$ are the regions of interest for the two images $\boldsymbol{X}$ and $\boldsymbol{Y}$.
The strength of MI as a similarity measure lies in the fact that no assumptions are made regarding the nature of the relation between the intensity values of the image, as long as such a relationship exists. The main drawback of MI comes from its probabilistic nature. The measure relies on an accurate estimate of the probability density of the image intensities. As a result, its effectiveness decreases significantly when working with small regions within the image, where there is not enough intensity samples to accurately estimate such densities. As a result, MI need to be calculated over a relatively large region of the image, which can diminish its

ability to handle small changes between the moving and fixed images.

The gradient magnitude of the images can also be used as well as the predefined metrics. In [Su 2013], SSD of both image intensities and gradients are used to register CT and CT-Fluoroscopy images. A robust patch-based correlation ratio was investigated in [Rivaz 2015] for the registration of 3D US and MR images for Image-Guided Neurosurgery. The Robust Patch Based Correlation Radio (RaPTOR) computes local correlation ratio values on small patches and adds them to form a global cost function. This metric is proved to be invariant to large amounts of spatial intensity inhomogeneity.

2.1.1.2 Feature-based methods

Geometric or feature-based methods focus on establishing correspondances between significant anatomical locations or salient landmarks. The selection of features to be considered is an important step, that should gather ease and accuracy of detection. The features can be markers attached to the patient's anatomy or intrinsic anatomical features. While extrinsic features are defined using artificial markers attached to the patient's body, this approach can be uncomfortable and not so feasible due to the elasticity and movement of some human organs. As a result, intrinsic features are preferred for more reliable and comfortable registration. Points, curves and surfaces are the most common features extracted.

The advantage of such methods is their ability to handle multimodal image registration where intensity values may not directly correspond. They can also provide improved robustness against noise and artifacts, as well as a wider capture range for deformations. However, the effectiveness of geometric methods heavily depends on the accurate detection and matching of these features, which can be challenging and computationally intensive, especially in complex or cluttered images.

The Iterative Closest Point (ICP) algorithm introduced by Besl and McKay in [Besl 1992] is commonly used to measure the distance between two curves, two surfaces or two volumes. In ICP, each point or a selected set of points in the point cloud of the moving image is matched with the closest point in the point cloud of the reference image. It minimizes the least square rigid transformation that relates these sets of points and iterates until convergence is achieved. It is a simple and fast method that requires an accurate initialization to perform effectively and not converge to a local minimum [Farnia 2012], [Moradi 2012], [Sinko 2018]. Yavariabdi et al. in [Yavariabdi 2013] proposed a novel variational one-step deformable ICP method to register transavaginal ultrasound (TVUS) and MR images. First, the data is manually segmented by an expert. Second, a deformable ICP method is used to compute a dense deformation field while establishing point correspondences automatically.

In [Singh 2008], Singh et al. introduced a manual registration technique for real-time transrectal ultrasonography (TRUS) and pre-acquired MR images to assist in prostate biopsies. They manually registered the MR image sequence to the ultrasound (US) images using custom software, enabling real-time navigation and

feedback. However, this method depends on the accuracy of selecting control points and ensuring their correspondence in both images. To reduce user interaction, [Kadoury 2012] also presented automatic and semi-automatic registration methods as alternatives to the manual approach.

Recently, deep learning has been suggested as a tool to build efficient registration methods. These methods can automatically learn and extract intricate features from images, eliminating the need for manual feature selection and reducing operator-dependent variability [Haskins 2020].

2.1.2 Transformation Model

Transformation in image registration is defined as the coordinate mapping from the reference image domain to the source image domain. Transformation model can be subdivided into rigid and non-rigid. Rigid transformations handle global rotations and translations, whereas non-rigid deformable models can generate local in-plane and out-of-plane deformations.

2.1.2.1 Rigid models

Rigid registration identifies the six degrees of freedom, mainly consisting of three rotational and three translational transformations, that map each point in the source image to its corresponding point in the reference image. These models are effective in applications where there is minimal to no change in the shape or position of the structure being analyzed. Rigid transformation models are the most common choice in the literature for 2D/3D registration, since they can deal with in-plane and out-of-plane rotations and translations [Meskine 2013]. Rigid transformations preserve the spatial distance between any consecutive points, they can be defined using different parametric forms, some of them being recalled below:
A translation, i.e., a constant displacement over space

$$x' = x + t_x, y' = y + t_y,$$

where x and y are the coordinate of a pixel in the moving image before registration, x' and y' are the same coordinates after registration and t_x and t_y are the translation parameters in the x and y axes, and/or a rotation:

$$x' = x\,\cos\theta + y\,\sin\theta; \quad y' = -x\,\sin\theta + y\,\cos\theta,$$

where θ defines the clockwise rotation. The number of degrees of freedom equals 3 for 2D images and 6 for 3D images.
An affine transformation can be decomposed into a linear transformation and a translation. In the 2D case, the affine model has six independent parameters:

$$x' = a_{11}x + a_{12}y + t_x; \quad y' = a_{21}x + a_{22}y + t_y$$

Such models maintain spatial relationships between points and asserts that lines that are parallel before transformation remain parallel after transformation.

2.1.2.2 Non-Rigid models

Non-rigid models are suitable for registration tasks where there is an anatomical variability across subjects or imaged body organs, which undergo soft-tissue type of deformation. Non-rigid registration requires determining a deformation field for the images. On one hand, a representation with higher degrees of freedom tends to be more expressive and may achieve better registration. On the other hand, higher degrees of freedom require determining a larger number of variables, leading to increased computational costs. A proper representation of the deformation field needs to be chosen to ensure sufficient expressiveness for a precise alignment without incurring excessive computational costs. These approaches can be classified into different categories according to i) The way in which the distance between the data and model sets is approximated (e.g., point wise, distance field) ii) The transformation used for mapping data set towards model set (e.g., Thin Plate Splines (TPS), Free Form Deformations (FFD), Laplacian deformation, etc).

Point wise position variables

A straightforward approach to defining a deformation field is to consider the point positions on the new deformed image as variables and compute them through optimization [Liao 2009], [Huang 2011], [Yamazaki 2013]. However, in practice, due to the physical behavior modeled by the deformation, a vertex often cannot move independently of other vertices. This can result in redundant degrees of freedom when using pointwise position variables. To address this, additional regularization is typically applied to the variables, such as local shape preservation [Liao 2009], [Huang 2011] and local similarity [Yamazaki 2013], to ensure more meaningful and coherent deformations.

Pointwise Affine Transformation

Instead of considering point positions as variables, some methods define an affine transformation (A_i, t_i) for each point on the source image

$$x_i' = A_i\, x_i + t_i,$$

where x_i is a point on the source image and x_i' is the point after deformation. $A_i \in \mathbb{R}^{3\times3}$ and $t_i \in \mathbb{R}^3$. Compared to pointwise position variables, this approach can better capture complex deformations, including local rotations. Based on the observation that the deformations in many non-rigid registration problems are locally rigid, affine transformations are often constrained to be close to rigid transformations. That is, the matrix $A_i \in \mathbb{R}^{3\times3}$ should be close to a rotation matrix [Amberg 2007].

Radial Basis Functions

Radial Basis Functions (RBFs) are a class of mathematical functions that depend only on the distance between points. They are typically used for interpolation and approximation of functions, particularly in scattered data interpolation and approximation problems. In their basic form, RBFs define a mapping from $\mathbb{R}^d$ to $\mathbb{R}$, where d is the dimension. A 2D RBF f is defined by a $\mathbb{R}^2 \to \mathbb{R}$ basis function ψ and a set of l centres q_k as:

$$f(x) = \sum_{k=1}^{l} w_k \psi(\|x - q_k\|).$$

It consists of a sum of l weighted terms with coefficients w_k of the basis function applied to the distance between x and the centre q_k. The basis function can have a localised influence (e.g. Gaussian) or a global behavior (e.g. TPS).
TPS are frequently used to generate a dense deformation field from a sparse set of control points. These methods involve a set of control points that can be located in arbitrary positions, which are usually obtained by detecting salient structures. TPS minimizes a bending energy based on this interpolation function, which gives a closed-form solution whose uniqueness is guaranteed in most cases. It can be decomposed into an affine and a local component [Chui 2003]:

$$x' = a_0 + a_1\, x + a_2\, y + \sum_{k=1}^{l} w_k \|x - q_k\|^2 \log(\|x - q_k\|),$$
$$y' = b_0 + b_1\, x + b_2\, y + \sum_{k=1}^{l} w_k \|x - q_k\|^2 \log(\|x - q_k\|).$$

In [Osechinskiy 2011], TPS are used to parametrize a smooth 3D deformation of a 2D slice. Control points are placed in a regular grid on the 2D image domain, and a 3D wrap is defined using three independent TPS functions.
In [Mitra 2010], an automatic deformable method for MR and US registration for prostate biopsy is proposed. The deformation model was based on TPS. One contribution was to estimate the spline knots using a triangulation technique with the axes of the segmented prostate in MR and TRUS images. A Shepard interpolation was used to overcome interpolation artefacts caused by the inverse TPS transform.

B-Splines

Cubic B-splines are a type of basis functions commonly used in interpolation and curve fitting. The term "B-spline" stands for "Basis spline," where "basis" refers to the piecewise polynomial functions that make up the spline. B-splines are defined recursively using a set of control points and a degree parameter. The B-splines used in medical image registration mainly consist of linear interpolating B-splines, convex nuclear B-splines, and cubic B-splines, which are the most commonly used. They

have a degree of 3, and provide a good balance between simplicity and accuracy.

Denote as $\Omega = \{(x, y) \mid 0 \leq x < N,\ 0 \leq y < M\}$ the domain of the image-plane and consider an $n_x \times n_y$ mesh of control points ϕ_{ij} with identical spacing. The cubic B-spline local transformation is defined as:

$$T_l(x, y) = \sum_{m=0}^{3} \sum_{n=0}^{3} B_m\Big(\frac{x}{n_x}\Big) B_l\Big(\frac{y}{n_y}\Big) \phi_{i_x+m, j_y+n}, \tag{2.4}$$

with $i_x = \lfloor\ x/n_x \rfloor - 1$, $j_y = \lfloor y/n_y \rfloor - 1$, and $\lfloor . \rfloor$ is used for the integer part. The cubic B-spline uniform functions are defined as:

$$\begin{aligned}
B_0(u) &= \frac{1}{6}(1-u)^3, \\
B_1(u) &= \frac{1}{6}(3u^3 - 6u^2 + 4), \\
B_2(u) &= \frac{1}{6}(-3u^3 + 3u^2 + 3u + 1), \\
B_3(u) &= \frac{1}{6}u^3.
\end{aligned}$$

After constructing the B-spline grid, it is used to deform the local neighborhood of every control point in the moving image. Once the moving image has been deformed, it is compared to the reference image using a similarity measure [Klein 2007].

Free-form deformation models

Free-form deformation (FFD) transforms are commonly employed in non-rigid image registration to model complex deformations of anatomical structures. They allow for local deformations by defining a grid of control points that can be moved to reshape and deform the image [Rueckert 1999]. In this model, the weighting functions are cubic B-splines. Unlike Thin-Plate Splines (TPS), where control points influence the entire domain during interpolation, the control points in this model have limited local support. They are uniformly distributed over the image domain in a grid-like form. In [Ferrante 2015], FFDs were used in a graph-based discrete optimization framework to perform slice-to-volume deformable mapping. In this model, a 2D grid-like graph simultaneously encodes the plane position (rigid body transformation) and the in-plane deformation of a slice with respect to its corresponding position at the 3D. The in-plane deformation is achieved through FFD interpolation, with the graph nodes serving as control points.

Fuerst et al. proposed an MR-US registration in [Fuerst 2014]. They used a free-form deformation based on cubic B-splines to build the transformation between the two images. They assumed that the US intensity is either correlated with the MR intensity or with the gradient of MR which leads to a first order polynomial function that relates the US image to the MR image and its gradient. Their approach demonstrated promising results in image-guided neurosurgery.

2.1.3 Optimization Algorithms

Optimization algorithms are used to select the optimal transformation to best minimize/maximize the defined matching criterion. These algorithms can be separated into two categories depending on the nature of the variables: continuous and discrete. The former optimizes a function through real values, while the latter extracts discrete variables from a discrete set.

2.1.3.1 Continuous Algorithms

Optimization method based on continuous variables essentially looks for the zeroes of the differential of an objective function, employing various parameter definitions for their optimization processes. For instance, when the stepsize remains consistent across iterations, each step aims to minimize the search direction. Typically, the search direction usually takes advantage of first-order or second-order specific information. A mathematical formulation of this strategy can be given by:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \alpha_t d_t,$$

where $\boldsymbol{\theta}$ contains the parameters of the transformation, t is the iteration number, α_t is the step size or gain factor and d_t is the search direction at iteration t. Based on the variations in search strategies, the commonly employed optimization methods for continuous variables include:

- Gradient descent method (GD): It is an iterative optimization algorithm used to minimize a function by iteratively moving towards the direction of the steepest descent, as indicated by the negative gradient of the function. In essence, it operates by calculating the gradient of the objective function at the current point and then adjusting the parameters in the direction that reduces the function value the most. The magnitude of the adjustment is determined by a predefined stepsize, referred to as the learning rate, which influences the convergence speed and stability of the algorithm. While GD is conceptually straightforward and computationally efficient, its performance can be sensitive to the choice of the learning rate, requiring careful tuning to ensure convergence to the global minimum in complex optimization landscapes. The FFD registration algorithm proposed by [Rueckert 1999] is GD-based.

- Conjugate Gradient Method (CG): In the Conjugate Gradient method, each search direction is conjugate to the previous ones, meaning they are orthogonal with respect to the problem's quadratic form. This property allows the method to converge faster than Gradient Descent, especially for problems with poorly conditioned or ill-conditioned Hessians. CG has overall faster convergence rate and improved performance. [Elen 2010] used CG to estimate rigid slice-to-volume mapping functions.

- Quasi-Newton: The Quasi-Newton method has a convergence rate that is higher than that of GD. In addition, compared to the CG method, Quasi-Newton algorithms use second-order information to speed up convergence. Other factors contributing to this quick convergence are their accumulation and the use of iterative information. The search direction is computed based on an estimation of the Hessian, using provided information from the previous iteration. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm is a common quasi-Newton method. [Hansen 2014] used the limited-memory version of it to solve multi-slice to volume registration.

- Nelder-Mead: Also known as the downhill simplex method, it is a direct search method that does not require the calculation of derivatives. It relies on a notion of simplex "a special polytope of $n + 1$ vertices in n dimensions" to explore the space of solutions. The algorithm operates by performing a series of transformations on the simplex, such as reflection, expansion, contraction, and shrinkage, guided by the evaluations of the objective function at the simplex vertices. Although easy to implement and apply to nondifferentiable functions or cases where the gradient is unknown, Nelder-Mead may not always converge to the global minimum and is sensitive to the initialisation [Dréo 2006], [Fei 2003].

- Powell: Powell's conjugate direction is a non-gradient method that does not require the function to be differentiable, and no derivatives are taken. It explores the search space by performing bi-directional searches along N different vectors, and iterating until no significant improvement is made. Powell's algorithm is known for its simplicity and robustness, making it suitable for a wide range of optimization problems. [Smolikova-Wachowiak 2005] used Powell's algorithm to register two-dimensional cardiac images to preprocedural three-dimensional images for interventional applications.

2.1.3.2 Discrete Algorithms

Discrete optimization deals with variables that take a finite set of distinct values. In image registration applications, discrete variable optimization essentially uses the Markov random field (MRFs) to conduct the optimization. Discrete variable optimization has the advantage of high calculation efficiency, and is currently used in medical image registration. Besides being inherently gradient-free compared to most continuous models, discrete optimization algorithms are not as prone to stuck in local minimas when the functions are not convex. Indeed, even complex functions can be optimized in the discrete scenario if the search neighborhood is large enough. Other methods include belief propagation methods [Yang 2010], graph-based methods, etc. In [So 2011], a graphcut-based method for non-rigid registration of brain MR images was used. The problem was formulated as a discrete labeling problem, and optimized by using α-expansions, which is a powerful combinatorial optimization tool, capable of yielding either a global minimum or a local minimum in a

strong sense.

2.1.4 US-MR registration: State of art

Up to this point, we have provided a general literature review on image registration, exploring various methods, techniques, and applications across different fields of medical imaging. In this section, we will narrow down our focus to the specific context of this study, which involves the registration of Ultrasound (US) and Magnetic Resonance (MR) images. We will present some state of the art algorithms that have been used for these specific modalities, i.e., 2D US and 3D MR image registration.

Atehortua et al. presented a 3D multimodal registration strategy to fuse 3D real-time echocardiography images with cardiac cine MRI images [Atehortúa 2019]. This alignment was performed in a saliency space, which was designed to maximize the similarity between the two imaging modalities. This fusion improves the quality of the available information. The used method performs in two steps: temporal and spatial registrations. A temporal alignment was firstly achieved by non linearly matching pairs of correspondences between the two modalities using a dynamic time warping. A temporal registration was then carried out by applying nonrigid transformations in a common saliency space where normalized cross correlation between temporal pairs of salient volumes was maximized.

Marami et al. presented in [Marami 2014] an automatic method for dynamically tracking the deformation of a soft tissue based on registering pre-operative three-dimensional (3D) MR images to intra-operative two-dimensional (2D) US images. The metric used is intensity-based, modality independant neighborhood descriptor (MIND), and the registration is based on an elastic deformation model correlating the images in the spatial and temporal domains.

Craene et al. in [De Craene 2004] studied a non-rigid MR/US registration algorithm based on a finite element elastic deformation model to capture the deformation in the liver ablation and the prostate biopsy. The similarity measure used in this study was a mix between mutual information and linear elastic energy. The optimization was performed using a new method referred to as Perturbation Stochastic Approximation (PSA). The use of volumetric meshes applied to the surfaces of organs to build the deformation showed promising results.

Kadoury et al. investigated in [Kadoury 2012] an approach to rigidly register the preoperative MR and the intraoperative US to allow pertinent needle placement during the ablation of liver metastases. Their contribution consists of training a deformable model for unsupervised segmentation of the liver in the MR image and extracting automatically the liver boundaries from the US image. The registration was performed using a weighted ICP algorithm with internal energy as a penalization.

An alternative approach leverages the Modality INdependent Descriptor (MIND) [Heinrich 2012] and its variant, the Self-Similarity Context (SSC) [Heinrich 2013]. Unlike methods based on global intensity relationships, these

techniques utilize differences in predefined neighborhood maps. The work in [Rivaz 2014] introduced the Self-Similarity α-MI (SeSaMI) metric as a similarity measure tailored for B-spline registration of MR and US images. This method demonstrated robustness to signal nonstationarity and intensity distortions, delivering promising results in image-guided neurosurgery.

Ferrante et al. [Ferrante 2017] presented a comprehensive survey of the literature about slice-to-volume registration, presenting a categorical study of different algorithms for medical imaging.

An iconic slice-to-volume registration was also studied in [Fei 2004] in the context of image guided surgeries. The idea is that low-resolution single photon emission computed tomography can be registered with a high-resolution MRI volume, which could be subsequently fused with live-time interventional MRI. A 3D mutual information registration method is used for the first step, and a robust slice to volume registration algorithm with special features for the latter.

2.1.5 Conclusion

In the first part of this chapter, we presented a comprehensive literature review of image registration models, highlighting their principal components such as matching criteria, transformation models, and optimization algorithms. We then discussed various methods specifically used for MR/US registration. This concluded with an overview of the state-of-the-art in MR and US registration. In the second part of this chapter, we introduce the proposed method for slice to volume registration. This method incorporates first a global rigid transformation characterized by rotation and translation parameters, then is combined with a local deformation based on B-spline functions.

2.2 Rigid slice to volume registration

This section investigates a registration model of 3D magnetic resonance (MR) and 2D ultrasound (US) images. The main objective of this work is to propose a rigid slice to volume registration algorithm inspired by [Porchetto 2017], and to use it as a prior step for US/MR image fusion. The algorithm of [Porchetto 2017] was pioneer in using discrete methods to solve the challenging slice-to-volume registration task, giving promising results when compared to a continuous approach. The results can be further refined using a continuous method, generating more accurate estimations. The accuracy of the proposed registration method is analyzed by means of quantitative and qualitative tests conducted on experimental phantom and realistic synthetic data generated from an in vivo MRI volume, with a specific attention to endometriosis treatment.

2.2.1 Slice-to-volume Model

The primary objective of our method is to achieve 3D/2D registration, commonly referred to as slice-to-volume mapping. A 2D US image is registered to a 3D MR volume, by searching the best matching MR slice, not necessarily following one of the three orthogonal directions of acquisition. This slice is defined by three rotation angles and three translation parameters. Its resemblance to the US image is based on a similarity measure, which has an important impact on the registration results.

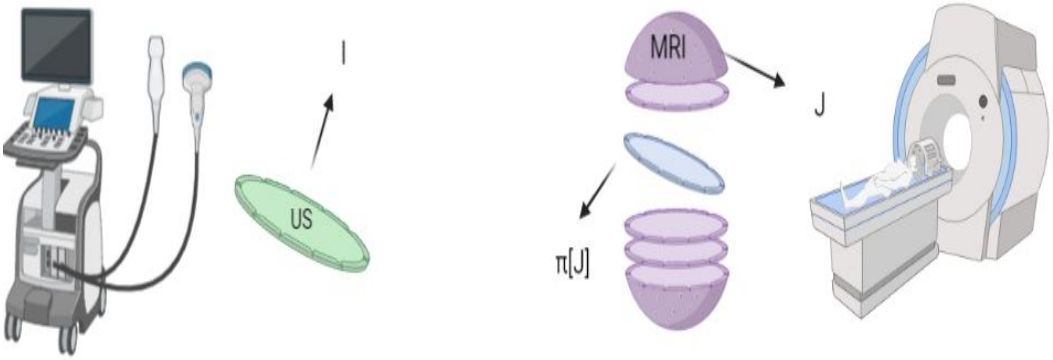


Figure 2.2: Schematic view of the proposed registration algorithm. I represents a 2D US image, related by a geometrical transformation π to a best-matching MRI slice $\pi[J]$.

Given an MR volume J and a 2D US image I, the goal of the method is to estimate the rigid transformation π defined by $\pi = (r_x, r_y, r_z, t_x, t_y, t_z)$, containing the rotation angles r and the translations t in the three spatial dimensions denoted by x, y and z, that better aligns I with a slice from J. This is achieved by solving the following optimization problem:

$$\widehat{\pi} = \underset{\pi}{\operatorname{argmin}}\ M(I, \pi[J]), \tag{2.5}$$

where $\pi[J]$ is a slice extracted from J specified by the rigid transformation π, as illustrated in Fig. 2.2 and M is the matching criteria that defines the dissimilarity between the 2D image I and the slice $\pi[J]$, which is inversely proportional to a similarity measure.
Depending on the characteristics of the images being registered, different matching criterias can be employed. In cases of monomodal images, where intensities exhibit linear correlation across both images, simple functions like the sum of absolute differences (SAD) or sum of squared differences (SSD) can effectively serve the purpose. However, in more complex scenarios such as multimodal registration, where the relationship between intensity values in the images is typically nonlinear, more sophisticated functions like mutual information (MI) are utilized.

The optimization problem associated with image registration is frequently tackled using continuous methods, both gradient-based and non-gradient-based. How-

ever, these methods are highly sensitive to initialization and may become trapped in local minima. To address this, the rigid slice-to-volume registration will be approached as a discrete labeling problem. This approach follows the discretization strategy proposed by [Zikic 2010].

2.2.2 Discrete labeling problem

Rigid slice-to-volume registration, along with numerous other problems in computer vision, can be modelled as a discrete labeling challenge within a pairwise Markov Random Field (MRF) framework. In formal terms, a discrete pairwise MRF constitutes an undirected graph $G = \langle V, E \rangle$, where each node $v_i \in V$, i = 1, ..., $|V|$, denotes a discrete variable. Two variables v_i and v_j are interdependent if there exists an edge $(v_i, v_j) \in E$ connecting their respective nodes. The potential values assignable to a discrete variable are constrained by the label space L. The objective of a discrete labelling problem within a pairwise MRF is to assign a label $l_i \in L$ to each $v_i \in V$, with the aim of minimizing the following energy function:

$$P(x; G; F) = \sum_{v_i \in V} g_i(l_i) + \sum_{(v_i, v_j) \in E} f_{ij}(l_i, l_j). \tag{2.6}$$

Here, $x = (l_1, ..., l_n)$ represents a labelling assigning a label l_i to each $v_i \in V$. The functions $G = g_i(.)$ correspond to the unary potentials associated with $v_i \in V$, while $F = f_{ij}(.,.)$ represent the pairwise potentials associated with edges $(v_i, v_j) \in E$. These functions yield scalar values upon assigning labels l_i to variables v_i. Since the optimization is cast as a minimization problem, potentials should assign lower values to labellings representing favorable solutions and higher values otherwise.

In equation (2.5), the aim is to explore the space of parameters π to identify the values giving the best matching. Since the problem is cast as discrete, we need a strategy to discretize the inherently continuous space of rigid transformations π. In their work [Zikic 2010], the authors demonstrated the feasibility of estimating linear transformations, particularly rigid body transformations, by solving a discrete and approximate version of this formulation.

Following their proposal, we model rigid slice-to-volume registration through a graph $G = \langle V, E \rangle$, associating every parameter of the rigid transformation $\pi = (r_x, r_y, r_z, t_x, t_y, t_z)$ to a variable $v_i \in V$, giving a total of 6 variables (nodes in the graph). G is a fully connected pairwise graph where $E = \{(v_i, v_j), \forall v_i \neq v_j\}$, meaning that all variables (parameters) depend on each other. In this strategy, every parameter v_i is updated through a discrete variation d_{l_i} associated with the label l_i. Figure 2.3 shows the topology of the MRF graph obtained for the registration problem, where the parameters of the transformation are represented by the nodes and the edges encode the image similarity costs associated with the variation of the labels of the adjacent nodes.

Each parameter of the transformation π will be constrained to a finite set and optimized in order to satisfy (2.5). More precisely, starting from an initial rigid transformation $\pi_0 = (r_x^0,\ r_y^0, r_z^0, t_x^0, t_y^0, t_z^0)$, the space of solutions is explored

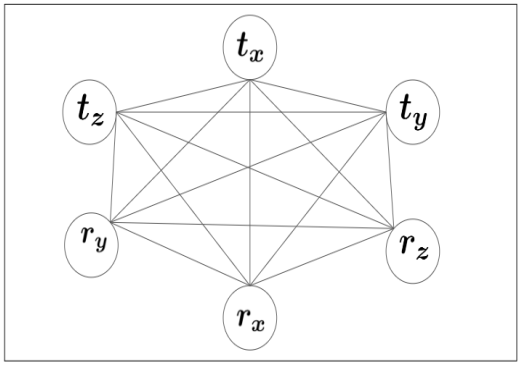


Figure 2.3: Topology of the 2D/3D registration MRF model. The parameters of the transformation are represented by the nodes. The edges encode the image similarity costs associated with the variation of the labels of the adjacent nodes.

by sampling discrete variations of π_0 to determine the transformation associated with the slice $\pi[J]$ best matching the image I through the similarity measure M. For a maximum size ω_i and a quantization factor k_i, the variations of the variable v_i associated with one component of π are constrained to be in the set $\{0, \pm\frac{\omega_i}{k_i}, \pm\frac{2\omega_i}{k_i}, ..., \pm\frac{k_i\omega_i}{k_i}\}$. Since the algorithm is guaranteed to reduce the energy in every step, we can assume that the new estimate is closer to the locally optimal solution, and reduce the search ranges for the parameters. Therefore, after each iteration, the maximum size ω_i is decreased by a factor α_i to allow a finer exploration of the search space. The total number of resulting values of v_i at iteration i is $l_i = 2k_i + 1$. Note that 0 is also included because the current parameter value can be preserved. For example, when the rotation variable r_x is considered, for $\omega_0 = 0.4$ and $k_0 = 2$, the search space for v_0 is $\{r_x^0, r_x^0 \pm 0.2, r_x^0 \pm 0.4\}$.

Since the number of possible solutions is exponential, it is not possible in a reasonable time to explore all the possible values of π. Instead, we adopt a pairwise approximation where only variations for pairs of variables are considered. These variations are encoded in the pairwise terms of the energy defined as $f_{ij}(l_i, l_j) = M(I, \pi_{l_i,l_j}[J])$. Here π_{l_i,l_j} denotes the updated version of π_0, where only v_i and v_j are modified according to the labels l_i and l_j, while the rest of the parameters remain fixed. Unary potentials g_i are not considered since we are only interested in the interaction between variables. Therefore, the discrete version of the optimization problem introduced in (2.5) becomes:

$$\hat{x} = \underset{x}{\operatorname{argmin}}\ P(x; F) = \underset{l_1,...,l_n}{\operatorname{argmin}} \sum_{(v_i,v_j)\in\epsilon} M(I, \pi_{l_i,l_j}[J]), \tag{2.7}$$

where the optimal label vector $\hat{x}$ represents the final rigid transformation $\hat{\pi}$ used to extract the solution slice $\hat{\pi}[J]$. Note that this pairwise model is clearly an approximation, since the real dependency between the parameters is not necessarily

pairwise. However, as stated in [Zikic 2010], similar approximations have shown to be good enough to estimate linear transformations, while making the problem tractable.

The key idea behind this approximation is related to the order of the model. The order of an MRF model is the maximum size of the involved cliques. It expresses the degree of conditional dependence of parameters on each other. Consider a simple first-order model: in this scenario, assigning a value to one parameter does not affect the labeling of other parameters within the MRF energy term. As a result, the labeling of individual parameters can be done independently. Going one step further and looking at a second-order model, we see that the choice of the label for one parameter will influence the choice for the labeling of all the parameters that form cliques together with it. Optimization methods for higher-order models have been proposed, but they are still not as general and efficient, and are complex to use. In [Zikic 2010], authors proposed an approximation to optimize high-order models by using second order cliques in the graph only. This means that the second-order cliques in the graph are used to encode a cost of a simultaneous variation of two parameters, while the other parameters are fixed to their current values. The MRF determines the parameters by optimizing the sum of the edge costs. This encodes the dependency of the similarity measure on the two respective parameters, while the dependency on all the other parameters is ignored. However, by simultaneously taking into account all possible combinations of parameter pairs, the overall energy term is constructed such that the selection of one parameter value depends on all the others.

2.2.3 Discrete method: FastPD

Discrete methods offer several advantages over continuous approaches in the context of image registration, as explained in Section 2.1.3. We propose to solve the discrete multi-labeling problem in (2.7) using FastPD [Komodakis 2007]. FastPD is a discrete optimization algorithm based on principles from linear programming and primal dual strategies, which generalizes the α-expansion method [Komodakis 2008]. It solves a series of max-flow min-cut problems on a graph. In that sense, it is similar to alpha-expansion, which also performs MAP inference on multi-label problems by solving successive binary max-flow min-cut problems. The main difference between these approaches is the construction of the graph where max-flow min-cut algorithm is applied. The α-expansion constructs the binary problem by restricting the label space, so that the only options for a given variable are to remain in its current assignment, or to take a label alpha (which varies in every iteration). Instead, FastPD constructs these binary problems by performing a Linear Programming Relaxation (LPR) of the integer program that represents the discrete MRF formulation. It builds upon principles drawn from the duality theory of linear programming, applying the well known primal-dual schema to the relaxed version of the MRF integer programming formulation. One of the main advantages of FastPD is its modular-

ity/scalability, since it deals with a much wider class of problems than α-expansion, being an order of magnitude faster while providing the same optimality guarantees when performing metric labeling.

2.2.4 Refinement Step

To enhance the accuracy and efficiency of the optimization process, the final discrete solution is used as initialization, and a continuous optimization strategy based on a simplex method (Nelder-mead) is finally run to further improve the result. The cost function to minimize is the same matching criteria M chosen before. This hybrid approach leverages the strengths of both discrete and continuous optimization techniques, aiming to strike a balance between computational efficiency and solution quality.

2.3 Experiments

In this section, we present the experimental evaluations conducted to validate the slice-to-volume 2D US and 3D MR images model. The primary application under investigation is the diagnosis of endometriosis. To assess the efficacy and accuracy of the proposed model, we initially validate it using an experimental phantom, for which both MRI and US acquisitions were performed. This phantom-based evaluation allowed us to simulate realistic imaging scenarios and assess the robustness of the chosen registration approach. Subsequently, our model was tested on a synthetic dataset. This dataset comprises a true volume MRI scan of a patient with endometriosis, from which we generate the corresponding observed MR and US images. For all the experiments, we use this fixed parameter setting: maximum translation size $\omega_{trans} = 7$, maximum rotation size $\omega_{rot} = 0.02$, quantization factor $k = 2$, and modification rate $\alpha = 0.8$.

2.3.1 Phantom Data

This section evaluates the results of the proposed registration model using a phantom dataset. The experimental phantom was designed to mimic uterus and endometrium responses to MR and US imaging. It was made of a beef steak on top of which was stuck a polyvinyl alcohol (PVA) phantom, using cyanoacrylate instant glue. The beef meat is composed of muscular tissues and its echogenicity and response to MR are similar to those of uterus tissue. The PVA phantom has roughly the same echogeneicity as the beef meat, but has different magnetic properties resulting in high contrast in the MR image. From this viewpoint, its properties are similar to those of endometrium. Finally, the glue between the two structures is visible on US image due to its high resolution and absent in MR image because of its limited resolution, mimicking the depth of penetration information, a crucial element for the surgery.

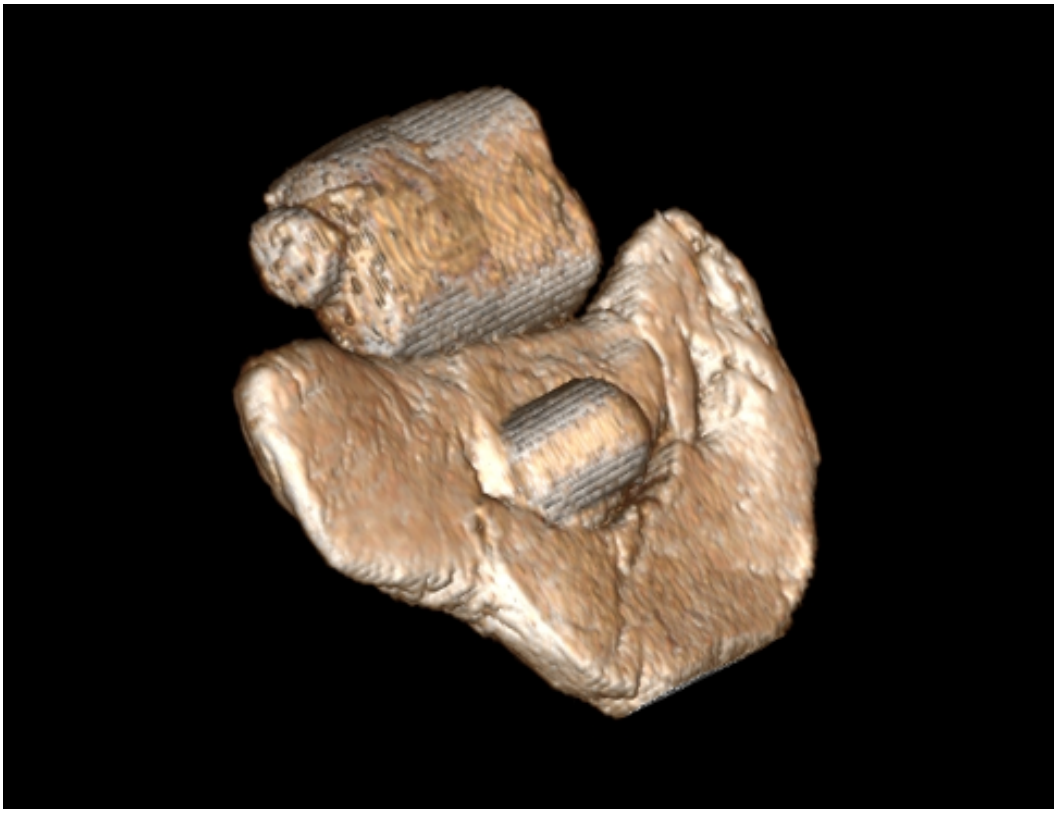
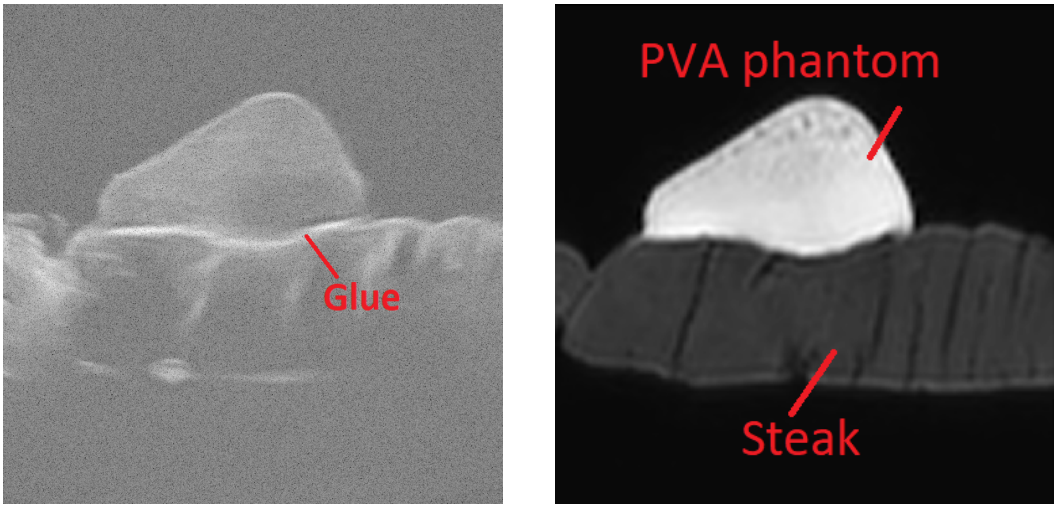


Figure 2.4: 3D MRI volume of the experimental phantom.

MRI acquisitions were performed using a 3T clinical imaging system (Philips Achieva dStream, Inserm/UPS UMR 1214, ToNIC Technical plateform, Toulouse, France). Axial fat-suppressed T1-weighted sequences (multishot mode; 4 mm slice thickness; voxel matrix $4 \times 1 \times 4$ mm) and axial, sagittal and coronal T2-weighted sequences (multishot mode; 2 mm slice thickness; voxel matrix $0.8 \times 2 \times 2$ mm) were acquired. For US image acquisition, the experimental phantom was immersed in a bucket full of water. A US examination was performed using a Voluson S10 system (General Electrics, USA). All images were acquired with a 10-MHz linear array transducer. More details about the experimental model design and image acquisition can be found in [Vidal 2019].

The dimensions of the images differ between the US and the MR modalities. Indeed, the US image has 600×600 pixels whereas the MRI volume has $320 \times 320 \times 90$ voxels. Given the wider field of view of the MRI, a manual cropping was done to adjust the MR volume to $100 \times 100 \times 90$, so as to ensure similar fields of view for both modalities. Finally, a despeckling of the US and a bicubic interpolation of the MR image were performed to ensure similar pixel dimensions.



(a) US image (b) Ground truth MR

Figure 2.5: Pair of MR slice and US image corresponding to manual 3D-2D registration.

For the 3D-2D registration task, the best matching MR slice from the volume corresponding to the 2D US (2.5(a)) image was manually identified with the help of an expert gynecologist surgeon. This slice will serve as the ground truth and is shown in Figure 2.5(b). The proposed registration algorithm was then tested using various initialization settings, representing MR slices with positions ranging from minor to more substantial deviations from the manual ground truth. The intensity-based metric used for performance evaluation in this example is the sum of squared differences (SSD). While this metric is not always the optimal choice for multimodal image alignment due to the differences in intensity distributions between MR and US, it is a suitable metric to explore in the case of the phantom since the images here exhibit close resemblance and share identical structures. Section 2.3.1.1 provides a comparative analysis of SSD with other commonly employed metrics, namely the Pearson Correlation Coefficient (CC) and the Mutual Information (MI).
The performance of the registration method was evaluated qualitatively through visual inspection of the registered image pairs, and quantitatively using the root mean square error (RMSE) defined as:

$$\text{RMSE} = \sqrt{\frac{1}{N}\|\hat{\boldsymbol{\pi}}[J] - \boldsymbol{x}_{\text{true}}\|_2^2},$$

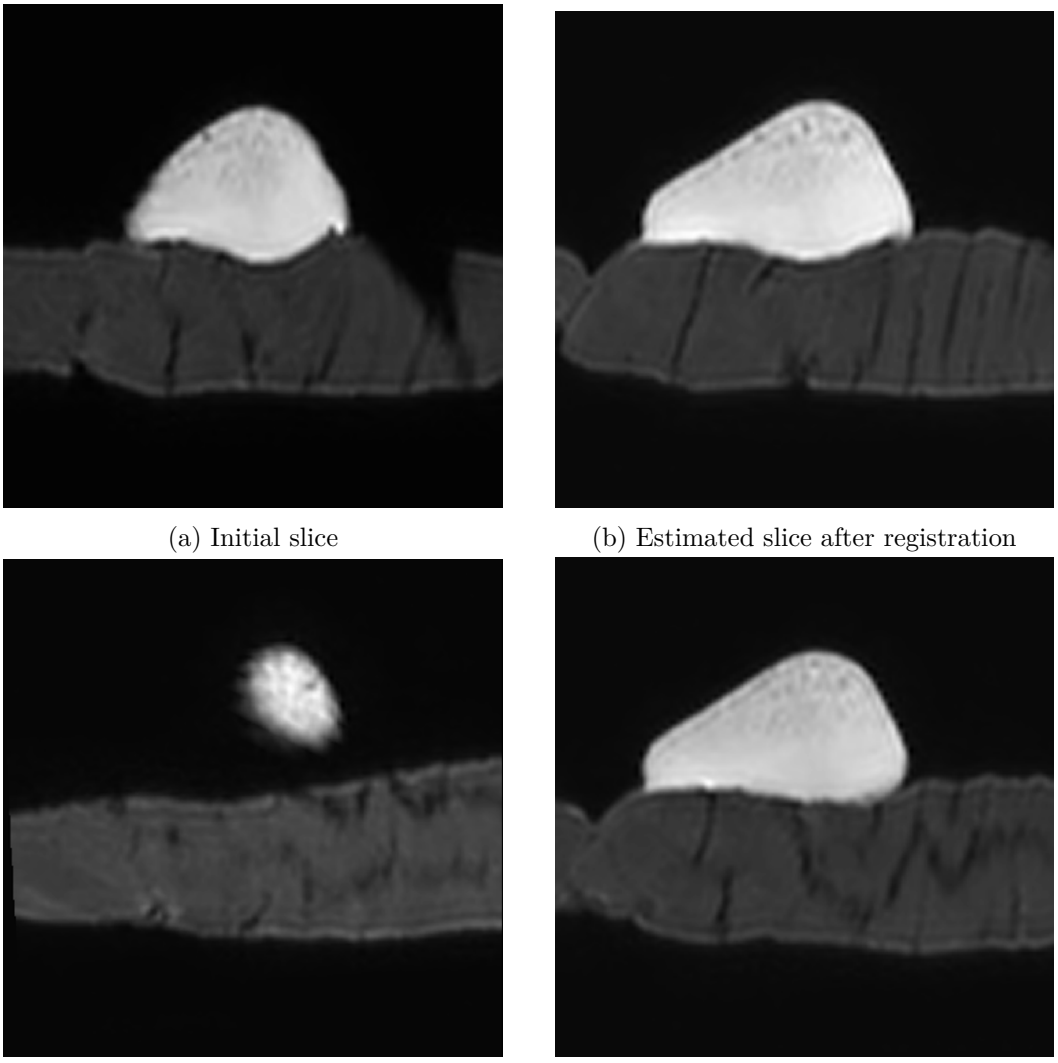
where $\boldsymbol{x}_{\text{true}}$ is the MR slice extracted manually to best match the US image, and $\hat{\boldsymbol{\pi}}[J]$ is the MR slice automatically estimated by the registration algorithm. The ultimate goal of this work is to create a reliable framework that produces a final informative image from the US and MRI. To this end, we will use an existing algorithm for 2D-2D MR/US image fusion [El Mansouri 2020], specifically designed for the same endometriosis diagnosis application. The rationale behind this ap-

proach is that an accurate registration should facilitate better fusion results, as misalignment in registration would inevitably lead to less accurate fusion results. Post-fusion evaluation will be conducted using the contrast-to-noise ratio (CNR) and the profile slope between two structures computed on the fused image. For two patches extracted from two different structures (PVA phantom and beef steak in this case), the CNR is defined as:

$$\mathrm{CNR} = \frac{|\mu_i - \mu_j|}{\sqrt{\sigma_i^2 + \sigma_j^2}},$$

where $\mu_i, \mu_j, \sigma_i^2, \sigma_j^2$ are the means and standard deviations of two blocks of pixels. Figs. 2.6 (a) and (c) show two initial slices given as input to the registration algorithm, corresponding to close and far locations from the ground truth. The registration results obtained for these initializations are displayed in Figs. 2.6 (b) and (d), showing that the estimated MR images are close to the ground truth of Fig. 2.5 (b) in both cases.

An example of the progression of the algorithm towards finding the final solution is shown in Fig. 2.7. Starting from the initial image depicted in Fig. 2.6 (a), we showcase the extracted slice after 1, 2, 3, 5, 30 and 60 iterations.



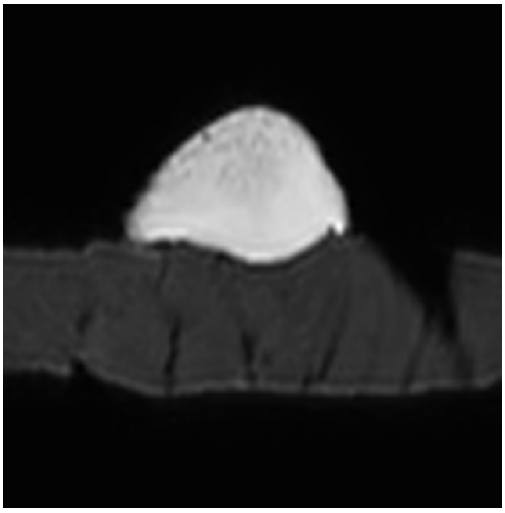
(a) Initial slice

(b) Estimated slice after registration

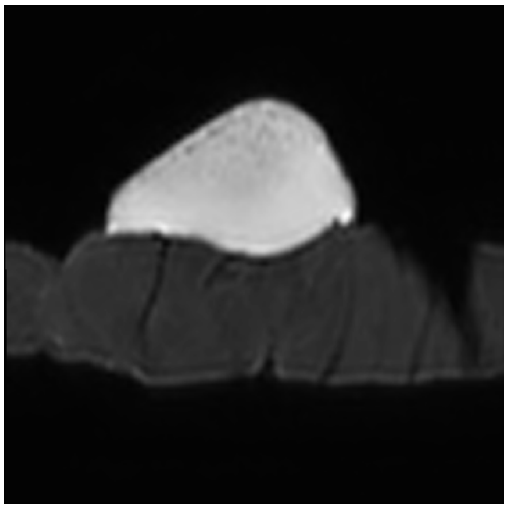
(c) Initial slice

(d) Estimated slice after registration

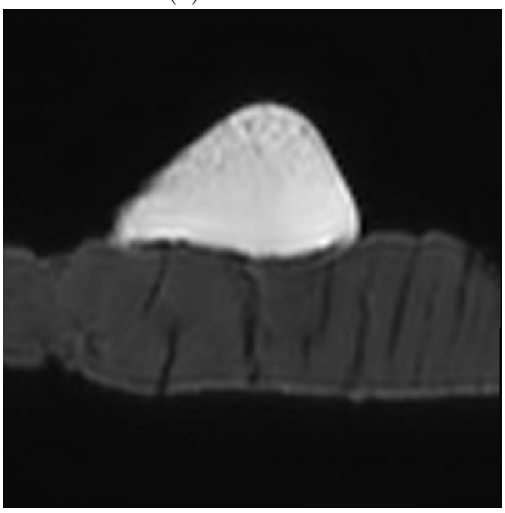
Figure 2.6: Image registration results obtained using an initialization slice with a small then an important deviation from the manual ground truth.



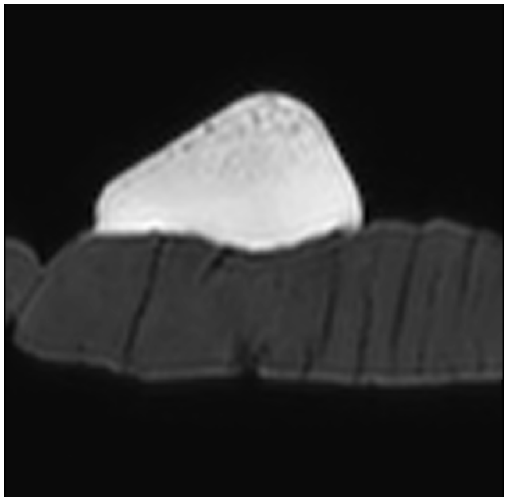
(a) Initial slice



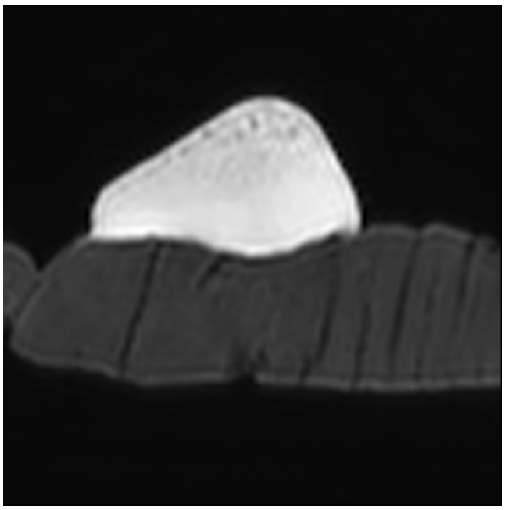
(b) Iteration 5



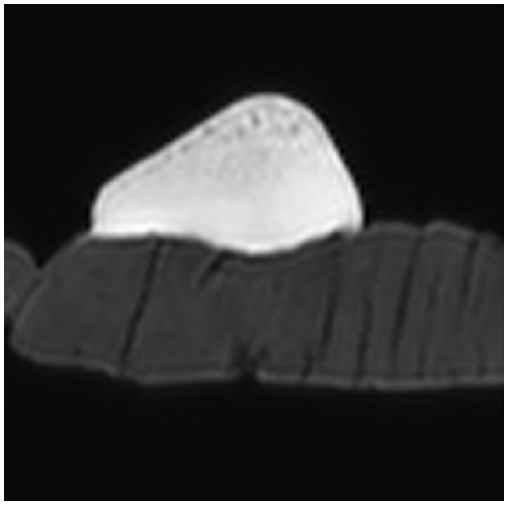
(c) Iteration 10



(d) Iteration 30



(e) Iteration 60



(f) Final result

Figure 2.7: Progression of the registration algorithm towards finding the final solution

In real life situations, the initialisation can be made using anatomical markers or based on acquisition knowledge, so that its deviation with respect to the MR slice best matching to the US image can be assumed to be small. Therefore, the following quantitative analysis focuses on the first set of images presented in Fig 2.6(a)-(b). The RMSE between the ground truth and estimated MR images is $7e^{-4}$, which is very promising. This result shows that using SSD as a similarity metric can be sufficient for this experimental case, due to the similarity between the images acquired on the phantom.

Fig. 2.8 shows the fused image (US image in Fig. 2.5(a) and extracted MR slice after registration in Fig. 2.6(b)). The fusion is based on an inverse problem, performing a super-resolution of the MR image and a denoising of the US image [El Mansouri 2020]. The fused image has a good contrast, comparable to the MRI, and a good spatial resolution similar to the US image. Moreover, the fused image can differentiate neighbouring structures and highlight small structures as the glue, contrary to the MR image. For example, the part between the PVA phantom and the beef meat is not distinguishable in the MRI, while it is clearly visible in the US and fused images. These results are confirmed in Table 2.1 presenting the CNR values and the slopes computed at the frontier between the steak and the glue. Note that the slopes are commonly used as an indicator of spatial resolution in US imaging.

	MRI	US	Fused image
CNR	54.21 dB	18.91 dB	43.17 dB
Slope	0.3×10^{-2}	1.5×10^{-2}	1.8×10^{-2}

Table 2.1: CNR and slope values after fusion.

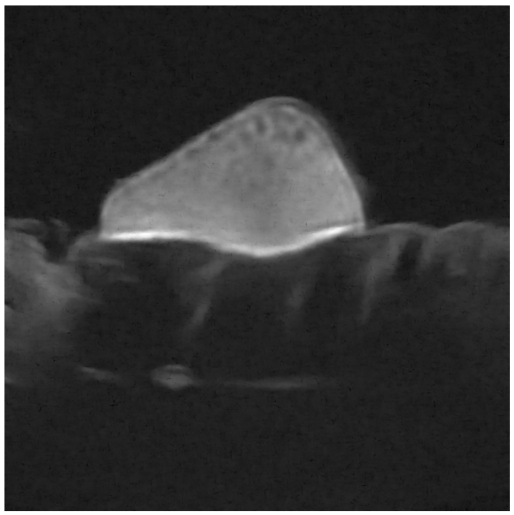
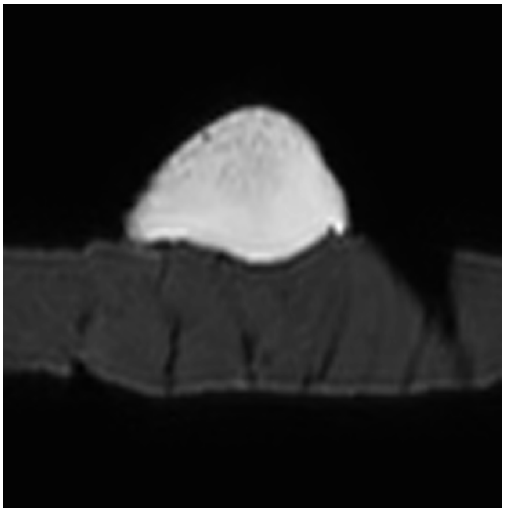


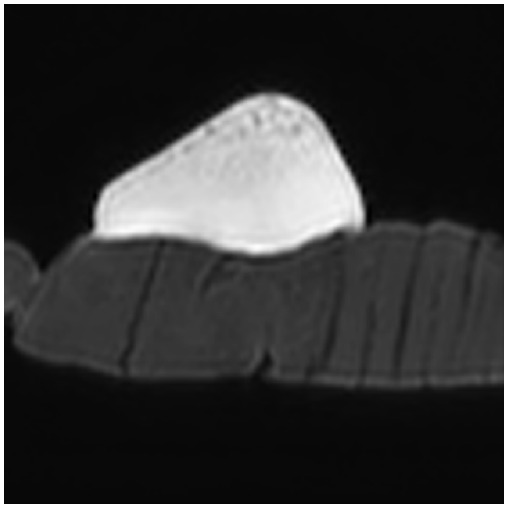
Figure 2.8: Fused image.

2.3.1.1 Comparaison between several intensity-metrics

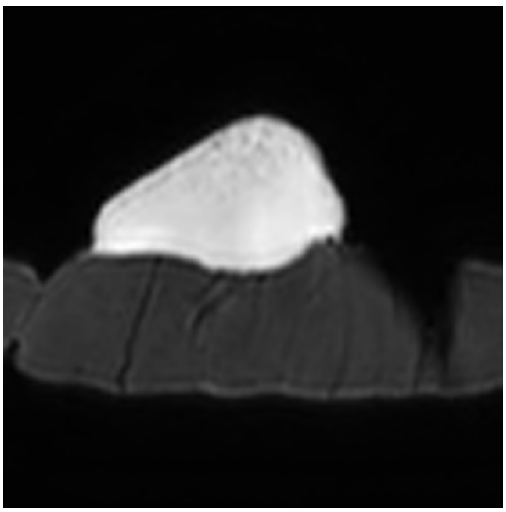
This section demonstrates the efficiency of the proposed MRI/US registration method using other intensity-based metrics. We use the same initial slices as before as input to the registration algorithm (Figs. 2.9(a) and 2.10(a)). Figs. 2.9 and 2.10(b,c) show the registration results obtained for different initializations, using the two tested similarity metrics. The results are visually close to the US image in Fig. 2.5 (a) for CC, but less accurate for MI as highlighted by the results in Fig 2.9(c) and 2.10(c), where a noticeable cut in the beef steak and a different PVA phantom shape is observed.



(a) Initial slice

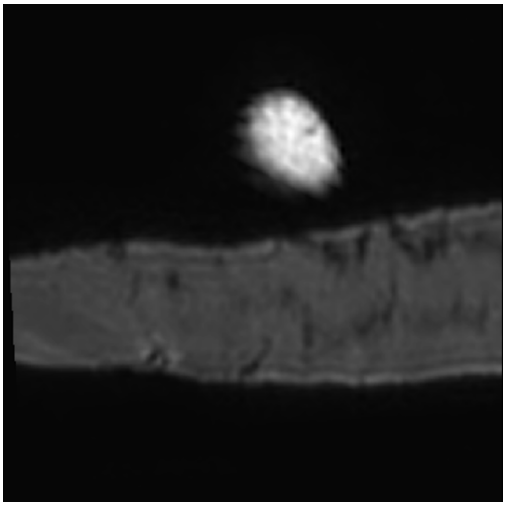


(b) Registration using CC

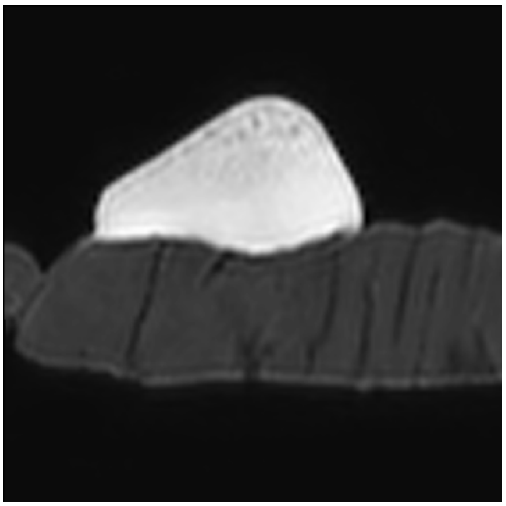


(c) Registration using MI

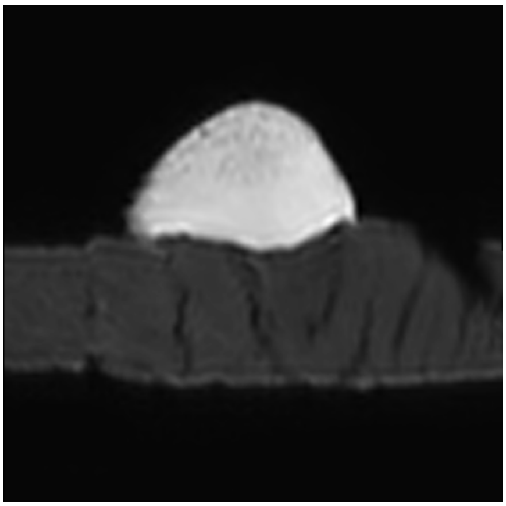
Figure 2.9: Image registration results obtained using an initialisation slice with a small deviation from the manual ground truth.



(a) Initial slice



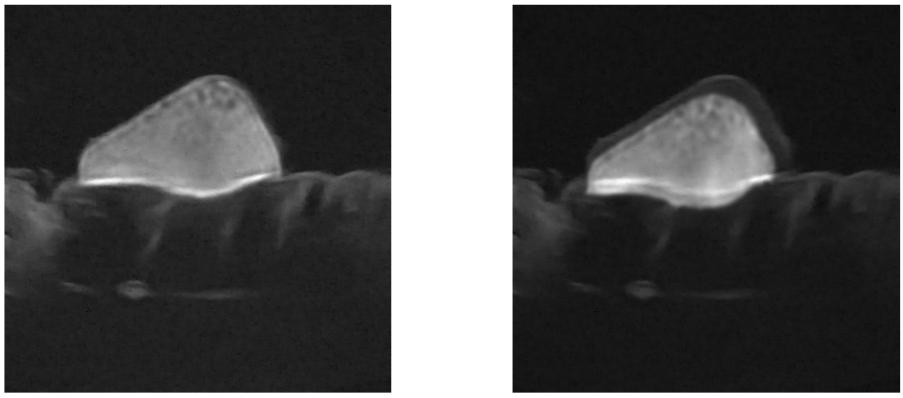
(b) Registration using CC



(c) Registration using MI

Figure 2.10: Image registration results obtained using an initialisation slice with an important deviation from the manual ground truth.

Fig 2.11 shows the fused images estimated from the US image and the MR slices estimated using both CC and MI. The fused images have a good contrast, comparable to the MRI, and a good spatial resolution similar to the US image. Note that the artifact visible in Fig 2.11(b) is caused by the misalignment induced by the registration error. These results are confirmed by the CNR and slope values in Tables 2.2 and 2.3. As we can observe, MI does not yield the best results compared to Cross-Correlation CC or SSD. This is because MI relies on an accurate estimate of the probability density of the image intensities. As a result, its effectiveness decreases significantly when the images have poor statistical consistency or lack clear structure. Examples of this include cases where there is overwhelming noise or conversely, when the area has very homogeneous intensities and provides very little information. In the phantom scenario, the images have black background with the beef steak in the middle, MI is focusing on the statistical dependence between pixel intensities and may not fully capture the spatial information or structural similarities between images.



(a) Fused image - CC (b) Fused image - MI

Figure 2.11: Image fusion results.

	MRI	US	Fused image
SSD	54.21 dB	18.91 dB	43.17 dB
CC	54.41 dB	18.91 dB	42.93 dB
MI	53.91 dB	18.91 dB	34.74 dB

Table 2.2: CNR values after fusion.

2.3.2 Synthetic data

This section presents the results of the proposed registration model using a synthetic data set with controlled ground truth. The simulations presented in this

	MRI	US	Fused image
SSD	0.3×10^{-2}	1.5×10^{-2}	1.8×10^{-2}
CC	0.3×10^{-2}	1.5×10^{-2}	2.1×10^{-2}
MI	3.4×10^{-4}	1.5×10^{-2}	1.66×10^{-2}

Table 2.3: Slope values after fusion.

section have been obtained using a real high-resolution MR image which was intentionally degraded to simulate an image resembling that obtained for endometriosis acquisitions. The original 3D high-resolution MRI volume in Fig. 2.12 captures detailed anatomical structures, including the uterus, bladder, and endometriosis lesions within the pelvic region.

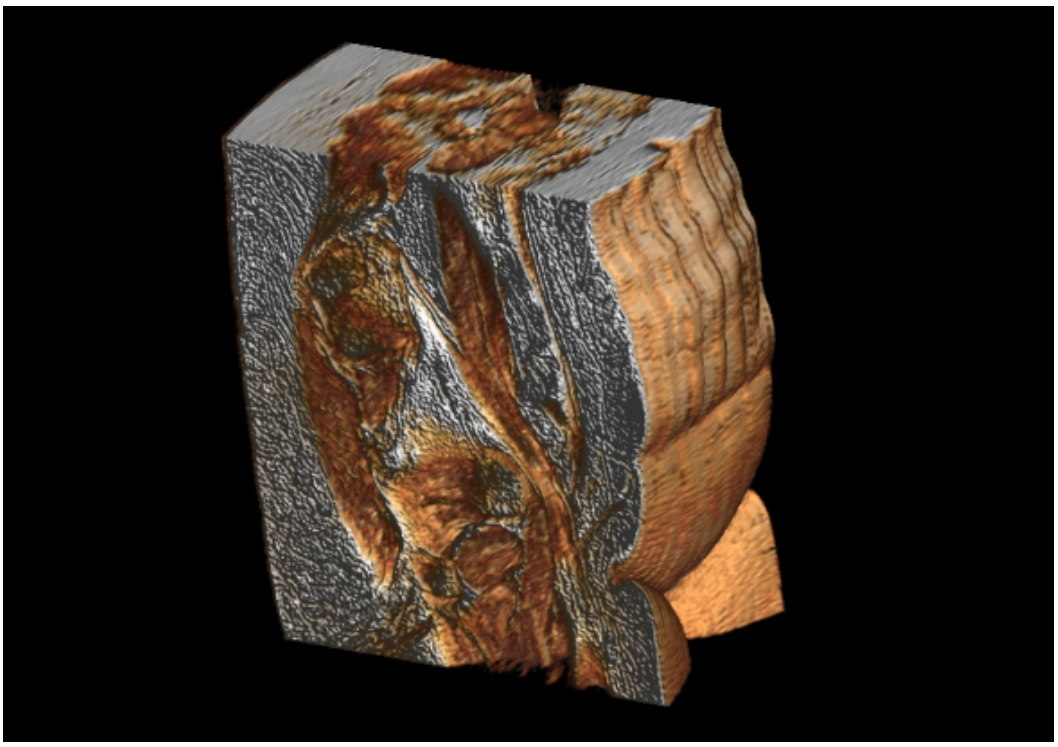


Figure 2.12: 3D MRI volume of a patient with endometriosis.

To create a degraded version of this high-resolution MRI, a combination of noise and blurring was applied. Specifically, the high-resolution volume was contaminated by an additive white Gaussian noise with a Signal-to-Noise Ratio (SNR) of 21.5 dB. Additionally, a 2D Gaussian filter with a standard deviation of $\sigma^2 = 4$ was used for the blurring. The ground truth high-resolution MRI image is shown in Figure 2.13 (a), whereas the initial slice extracted from the degraded volume and used as an input to the registration algorithm is displayed in Fig. 2.13 (b).

To simulate the US image, a third-order polynomial transformation, as in [El Mansouri 2020], was applied to the clean high-resolution MRI image. This

transformation was corrupted by a log- Rayleigh additive noise to simulate the typical speckle observed in clinical US imaging, yielding the image displayed in Fig. 2.13 (c), with an SNR of 8 dB. The intensity-based metric used in this section is the mutual information. Given the initial slice in Fig. 2.13 (b), Fig. 2.13(d) shows the estimated registered image, with an RMSE between the ground truth and estimated MR images equal to RMSE = 0.02. The fused image obtained using the two registered US and MR images is displayed in Fig. 2.13(e). This image provides a good compromise between the US and MR data. Specifically, the fused image is less affected by US speckle and MRI blur, provides well-defined contours and good contrast compared to the native MR and US images.
In addition to the visual inspection of the different images, CNR was used to evaluate the contrast between two different structures of the images. The two regions considered here are extracted from the uterus and the bladder. The CNR values for the MR and the US images are respectively 41 dB and 19.17 dB. The final fused image had a CNR of 39.01 dB, which clearly demonstrates that the fusion process improves the contrast in the images compared to the US image.

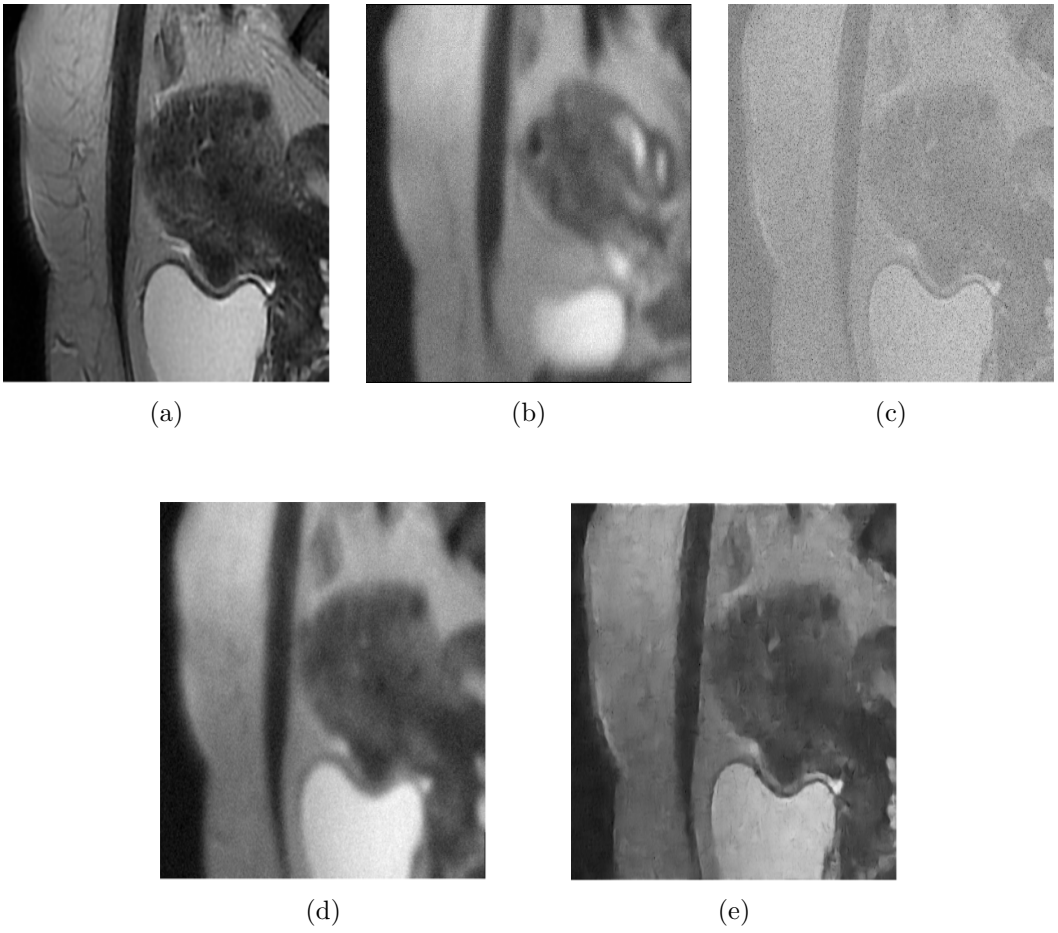


Figure 2.13: (a) True high resolution MR image (Ground truth). (b) Initial MRI slice (MR low-resolution and blurred image). (c) US image (polynomial function of the noiseless MRI with additive log-Rayleigh noise). (d) Extracted slice after registration. (e) Fused image.

2.4 Non-Rigid Registration

The 2D/3D rigid registration method presented earlier incorporates a global rigid transformation characterized by rotation and translation parameters to align the MR slice with the corresponding US image effectively. Although the experiments conducted demonstrated the efficiency and reliability of the approach, real-world scenarios present challenges since US and MRI scans are conducted under different conditions. For instance, in pelvic US, patients are advised to drink plenty of water to fill the bladder, facilitating better imaging by moving the uterus upwards and pushing the bowel away. Conversely, for MRI exams, patients are instructed to fast for at least four hours prior to the exam, resulting in a different bladder shape in the images. These variations generate local deformations that are not captured by

rigid registration. A non-rigid approach is therefore required to achieve accurate alignment in a real data scenario. We propose to use a B-spline deformation mesh. The basic principle of the method is to express the image deformation field using control points defining the local movement of their neighboring coordinates. The motion field of the points between the control points is then estimated using cubic B-spline functions.

Let $\Omega = \{(x, y) \mid 0 \leq x < N,\ 0 \leq y < M\}$ be the domain of the image-plan and consider an $n_x \times n_y$ mesh of control points ϕ_{ij}. The local B-spline transformation is defined by:

$$T_l(x, y) = \sum_{m=0}^{3} \sum_{n=0}^{3} B_m\Big(\frac{x}{n_x}\Big) B_l\Big(\frac{y}{n_y}\Big) \phi_{i_x+m, j_y+n}, \tag{2.8}$$

with $i_x = \lfloor\, x/n_x \rfloor - 1$, $j_y = \lfloor y/n_y \rfloor - 1$, where $\lfloor . \rfloor$ denotes the entire part. The uniform cubic B-spline functions considered in this study are defined as follows:

$$\begin{aligned}
B_0(u) &= \frac{1}{6}(1-u)^3, \\
B_1(u) &= \frac{1}{6}(3u^3 - 6u^2 + 4), \\
B_2(u) &= \frac{1}{6}(-3u^3 + 3u^2 + 3u + 1), \\
B_3(u) &= \frac{1}{6}u^3.
\end{aligned}$$

Once the grid of the B-spline has been constructed, it is used to distort the local neighborhood of each control point in the MRI image. The best deformation provides the distorted MRI image closest to the reference image (here, the US image), i.e., which minimizes the chosen similarity measure. Note that for each rigid transformation π_i estimated at iteration i using the rigid registration algorithm, an optimization of the parameters of the B-spline functions is carried out generating a new image $\pi_i'[J]$. Instead of seeking to minimize/maximize the (dis)similarity measure between the US image I and $\pi_i[J]$, the optimization will be carried out between I and $\pi_i'[J]$.

2.4.1 Real Data

This section considers experiments using real data to evaluate the non-rigid registration approach. For this, we use the previously mentioned real MRI volume along with a real abdominal US, moving away from the simulated US used in our earlier tests. This shift to authentic data allows us to assess the performance and robustness of the proposed registration method in more realistic clinical scenarios.

US imaging was performed using a Voluson P6 imaging system (GE Healthcare Austria GmbH & Co OG). For the MRI, acquisitions were performed using a 1.5T clinical imaging system (GE Medical Systems Signa). Axial, sagittal, and coronal T2-weighted sequences were obtained with a 4 mm slice thickness and 5 mm spacing between slices, ensuring adequate coverage of the pelvic region. The images were

acquired using a 2D acquisition mode with a repetition time (TR) of 3586 ms and an echo time (TE) of 138.0400 ms. The echo train length was set to 34, enhancing the quality of the T2-weighted sequences. The imaging frequency was 63.9053 MHz, targeting the hydrogen nuclei (1H) for optimal signal acquisition.

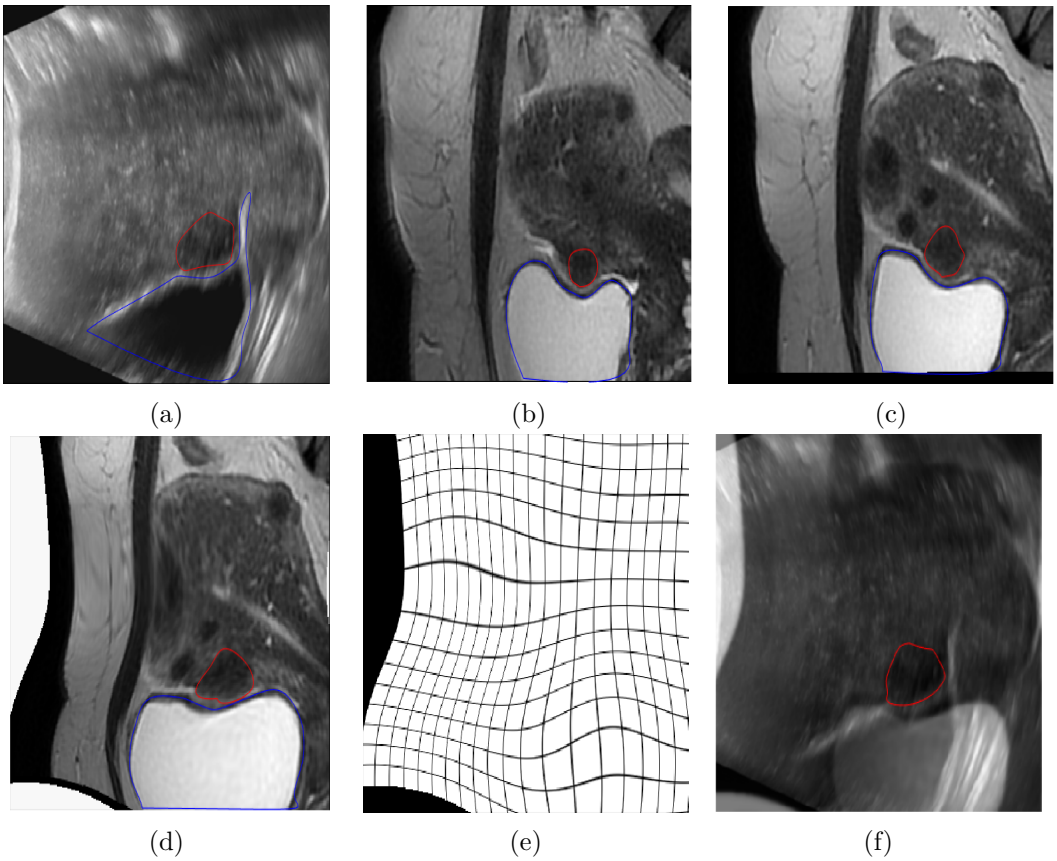


Figure 2.14: (a) US image (Endometriosis lesion - Red, Bladder - Blue). (b) Initial MR slice. (c) Estimated image after rigid registration. (d) Estimated image after non-rigid registration. (e) Deformation grid. (f) Fused image.

The sizes of the acquired images are 400×300 for the US image and $512 \times 512 \times 24$ for the MRI volume. As the field of view of the MRI image is wider than that of the US image, the volume was manually cut to have $220 \times 270 \times 24$ voxels, to ensure similar fields of view for both modalities. Finally, bicubic interpolation of the MRI image was performed to ensure the same pixel size in the MRI and US images. MI is used as a similarity metric for the registration since it measures the statistical dependency between the images and has proven to be a more reliable and versatile metric for image registration in multi-modal real data imaging.

Figures 2.14(a) and 2.14(b) show the observed US image and the MRI slice chosen as initialization of the 3D-2D registration code. Figures 2.14(c) and 2.14(d) show the results of the 2D-3D rigid registration, then the registration after taking

into account non-rigid deformations. It can be seen that the shapes of the bladder and endometriosis lesion in image 2.14(d) match those in image 2.14(a) more closely than in 2.14(c). The deformation of the grid carried out is shown in Fig. 2.14(e). The MI values in Table 2.4 confirm the benefit of the adjustment carried out. Finally, we also show the result of the 2D-2D fusion of US and MRI images in figure 2.14(f). This image brings together information from both modalities, allowing good localization of the endometriosis lesion (circled in red) in relation to the bladder, using US, and good contrast of the fibroid using MRI.

	Mutual Information	
	US vs Initial MR slice	US vs Estimated slice
(Fig. 2.14)	0.46	0.56

Table 2.4: Mutual information values between the US image and the initial MRI image or the final MRI image after registration.

2.5 Conclusion

The chapter focused on medical image registration, particularly tailored for our specific application, which is endometriosis diagnosis and surgery. We introduced various algorithms and components of a registration model, alongside state-of-the-art models. Subsequently, we presented a slice-to-volume framework based on a discrete labeling problem solved using a discrete optimization algorithm and further refined using a continuous optimization algorithm. Different similarity metrics were considered depending on the datasets used. Moreover, the proposed method was enhanced by incorporating non-rigid deformations to address challenges like bladder deformations. A natural progression of this multimodal registration work is to combine it with multimodal image fusion, which will bring us closer to the ultimate goal of this study: transforming an initial set of 3D MRI and 2D US images into a final informative image, which is the aim of the next chapter.

Chapter 3

2D/2D MR-US fusion: A non-parametric approach

Contents

Image fusion refers to the process of combining multiple images from different sources, modalities, or time points to create a single composite image that contains more information than any of the individual images alone. In this chapter, we will first present state-of-the-art models and methods used in image fusion. This will provide a comprehensive overview of how image fusion is approached in the literature, highlighting its advantages, disadvantages, and the challenges encountered in various applications. Then, we introduce the proposed fusion method for 2D Ultrasound (US) and 2D Magnetic Resonance (MR) images. The goal is to create a hybrid image (referred to as MARIUS image, for MAgnetic Resonance Imaging & UltraSound) gathering the advantages of both modalities. This method is built upon the idea developed in [El Mansouri 2020], which focuses on constructing two observation models associated with US and MR images to exploit the complementarity of both images in terms of resolution and contrast. For the MR image, a linear model formed by blurring and downsampling operators was proposed, motivated by its good performance for super-resolution [Gholipour 2010, Greenspan 2002, Manjón 2010]. Meanwhile, a denoising model was employed for the US image in order to mitigate the effect of speckle noise

[Aysal 2007a, Gupta 2005]. Since US and MR modalities are different by their nature and content, there is no simple correspondence between the gray levels of these images. We propose to use a non-parametric transformation defined using the theory of reproducing kernels and to evaluate its benefit when compared to the polynomial transformation investigated in [El Mansouri 2020]. The proposed fusion method is evaluated on an experimental phantom and synthetic data generated from an in vivo MRI volume, with a specific attention to endometriosis treatment.

3.1 Literature review

Image fusion refers to assembling all the important information from multiple images and including them in fewer images, e.g., in a single image. The purpose of image fusion is to build an enhanced image that is informative, comprehensible and accurate for the desired application [Amin-Naji 2018]. For instance, in remote sensing, combining images from optical and radar sensors can provide detailed and accurate terrain analysis [Zou 2022] [Joshi 2016]. Similarly, in surveillance, fusing data from infrared and visible light cameras can significantly improve target detection and identification in diverse lighting conditions [Kaur 2022]. In the field of environmental monitoring, merging satellite images with different spectral properties aids in better tracking and understanding of climate changes [Byun 2015], forest monitoring, land use [Sanli 2017], coastal zone evolution [Yang 2012] and natural disasters [Tlig 2022]. In medical imaging, image fusion is becoming increasingly common for the study of various pathologies [Muzammil 2020], diagnosis accuracy [Haribabu 2023], treatment planning [Krempien 2003] [Ge 2008] and monitoring [Giesel 2009]. By combining images from different modalities, clinicians can leverage the unique strengths of each modality to gain a more comprehensive understanding of a patient's condition [James 2014]. A large and growing body of literature has investigated various techniques to address these challenges. Researchers have explored different methods to fuse images, while maintaining the necessary information from all of them [Meher 2019, James 2014].

Fusion methods for medical images have their own advantages and disadvantages. They offer a comprehensive view of the patient's condition due to the combination of different modalities. This highlights specific features such as tumors or lesions, e.g., PET scans show metabolic changes that occur at a cellular level in an organ or tissue, which is where diseases can first be detected [Liao 2012]. MRIs cannot view the tissue at this scale, but rather are useful for detecting larger changes in the organ or tissues [Lladó 2012]. Fusing these modalities can lead to more accurate diagnoses as doctors can see different aspects of the anatomy and pathology in a single image [Bhavana 2015]. Fusion methods can also reduce the need for multiple diagnostic tests by providing sufficient information in one go, saving time and reducing patient discomfort and exposure to radiation [Goreczny 2017]. However, image fusion techniques can be sensitive to noise and introduce additional artifacts or er-

rors, particularly if the images are not perfectly aligned, or if the algorithms used are not optimal. The process of fusion can also be complex and require advanced software and equipment [Khan 2020]. This can lead to higher costs, which may not be affordable or available in all medical facilities, especially with the absence of a generic method applicable to all modalities [Tan 2020].

Image fusion models can be classified into three main categories: pixel level, feature level and decision level. Each of these approaches offers distinct advantages and methods for combining information from multiple images to enhance the final output:

- Pixel level techniques for image fusion involve directly integrating the raw data from input images to produce a fused image. This approach works by combining pixel intensities from different images, which allows for the preservation of detailed information from each source [Li 2017].
- Feature level techniques, on the other hand, focus on the extraction and integration of salient features from the input images. These features can include pixel intensities, textures, edges, and other significant details that are then compounded to create a richer, more informative merged image [Calhoun 2009]. These techniques are based on the enhancement of specific attributes in the images. They are useful and employed in various applications such as image recognition, where distinguishing different objects or patterns is crucial.
- Decision level fusion techniques take a different approach by processing input images separately to extract relevant information before combining them. As an example, the images of interest can be segmented into homogeneous regions and this segmentation can be used to define a fusion model [Liu 2017].

3.1.1 Image fusion techniques

The sharp increase in scientific papers on fusion techniques is driven by the demand for high-accuracy models at a low cost in various domains. There exist a lot of techniques on image fusion, each offering advantages and limitations and dealing with different modalities. Figure 3.1 shows a classification of these methods. In this section, we will introduce some of the most commonly employed fusion techniques [Kaur 2021]. Multimodal fusion methods typically operate in three main steps. Firstly, the input images are decomposed into sub-images or representations that capture specific aspects of the data from each modality. Secondly, fusion rules are applied to combine the information from these sub-images. Finally, the fused image is formed by using appropriate reconstruction algorithms.

3.1.1.1 Image decomposition and reconstruction

Numerous decomposition and reconstruction algorithms were proposed in the literature. Here, we present two fundamental methodologies commonly employed in

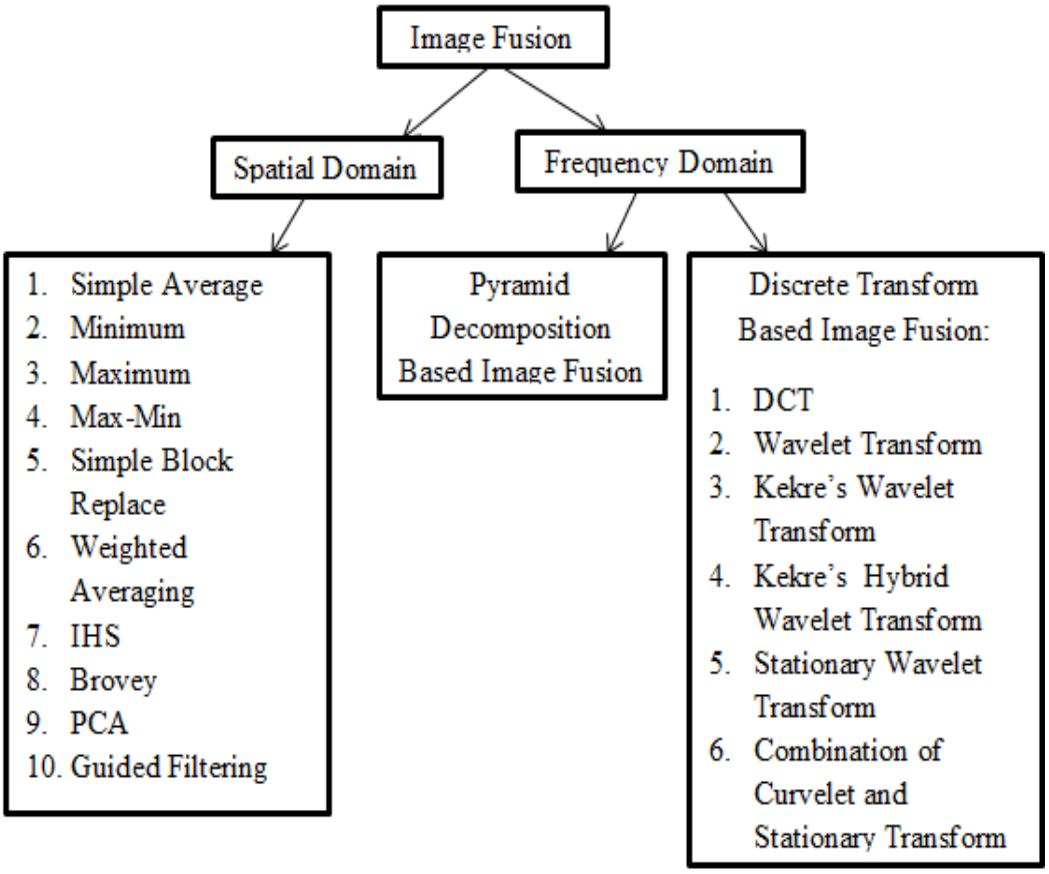


Figure 3.1: Image fusion techniques "Figure extracted from [Mishra 2015]".

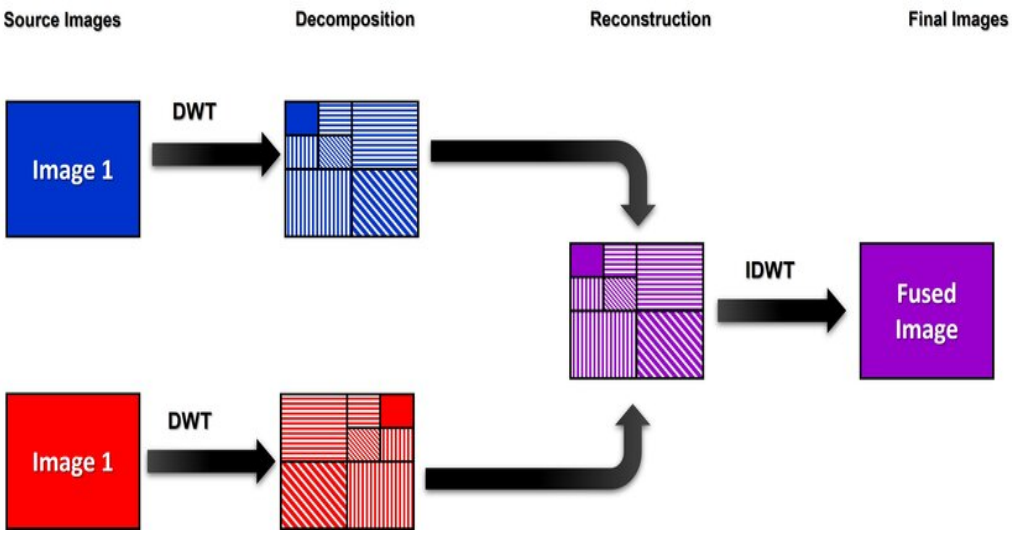


Figure 3.2: DWT-based Image fusion steps "Figure extracted from [TRIVEDI 2023]".

medical image fusion: discrete wavelet transform and Laplacian pyramid transform.

Discrete Wavelet Transform (DWT): The discrete wavelet transform (DWT) is a powerful tool for image fusion [Tao 2011, Maranur 2018]. In DWT-based image fusion, the source images are first transformed into the wavelet domain, where each image is represented by its wavelet coefficients. These coefficients capture both spatial and frequency information, which is crucial for preserving important features from the original images. The transformation process applies wavelet filters to the images, which results in the decomposition into approximation (low-frequency) and detail (high-frequency) subbands at multiple resolution levels. The approximation subband contains the coarse information, while the detail subbands capture edges and finer details. This filtering and downsampling procedure is repeated according to the desired level of decomposition. Typically, higher resolution images necessitate a greater level of decomposition compared to lower resolution images. Figure 3.2 shows the basic steps in image fusion process using DWTs. Once the images have been decomposed, the fusion process begins by combining the corresponding wavelet coefficients from each source image. Different fusion rules can be applied to select and merge these coefficients, such as choosing the maximum absolute value, averaging, or using more sophisticated adaptive techniques that consider the local context of the coefficients. The choice of fusion rule significantly impacts the quality of the fused image, as it determines how well the important features from each source image are preserved and integrated. After fusing the coefficients, an inverse discrete wavelet transform (IDWT) is applied to reconstruct the fused image from the combined coefficients.

A DWT-based fusion method was studied in [Reddy 2014] where a DWT architecture was considered to fuse images collected from various angles using

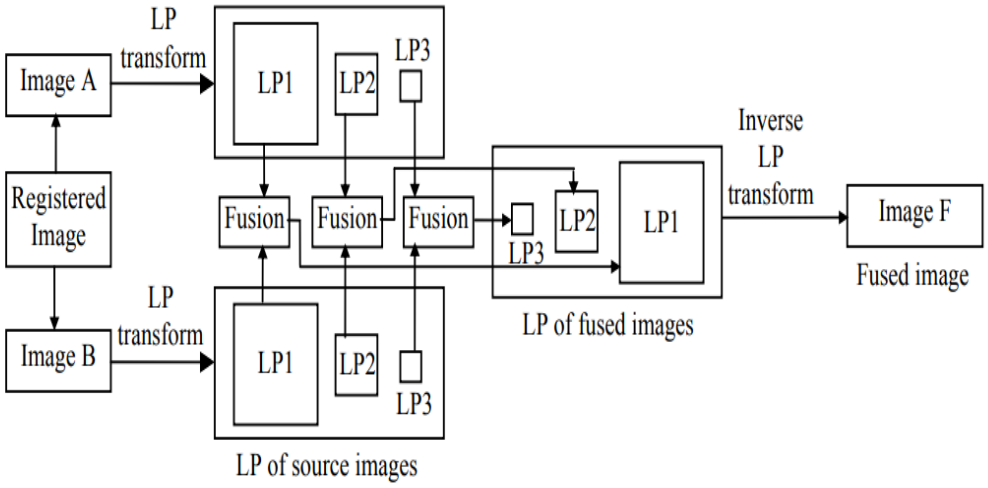


Figure 3.3: A fusion framework based on the Laplacian transform "Figure extracted from [Sabre 2018]".

Micro Air Vehicles. In [Garg 2005], a region level image fusion technique, has been implemented and analyzed for MR T1 and T2 images. The fusion strategy uses multi-level decomposition of the images obtained using the wavelet transform. Loza et. al. proposed a new methodology for multimodal image fusion based on non-Gaussian statistical modelling of wavelet coefficients, with a special emphasis on noisy images [Loza 2010].

Laplacian Pyramid Fusion: This technique utilizes a Gaussian filter as a blur operator and calculates the difference between images at adjacent pyramid levels to reconstruct the high-resolution image. Laplacian pyramid fusion consists of three phases: image decomposition, image fusion using precise fusion rules, and reconstruction of the fused image using the inverse Laplacian pyramid transform. Figure 3.3 shows the fusion framework based on the Laplacian Pyramid. This approach is widely used for fusing different imaging modalities, such as MR and CT scans, MR and SPECT scans, and MR and PET scans [Li 2018, Sahu 2014]. However, this method can be computationally expensive in term of data storage, particularly when processing large images. Saleem et al. proposed a contrast enhancement technique to improve the visibility of image details without introducing unrealistic visual appearances or unwanted artefacts. Their technique is based on a multi-resolution method using a Laplacian pyramid decomposition to account for the multi-channel properties of the human visual system [Saleem 2012]. Finally, we mention here other decomposition and reconstruction algorithms based on the Kekre's wavelet transform [Dhannawat 2013] or curvelets [Zhao 2023].

Deep learning: Recently, fusion methods based on deep learning techniques have emerged as an active field of research [Li 2021] [Zhang 2021]. The ability

of neural networks to predict, analyze and infer information from a given data without going through a rigorous mathematical solution makes them attractive to use for medical image processing, as the nature of variability between the images is subjected to change every time a new modality is used. Training neural networks to adapt to these changes enable several applications for image fusion. Convolutional neural network (CNN), which is a category of neural networks, usually is superior to traditional manual feature extractors in feature extraction. In [Zhu 2022], a general CNN framework for image fusion, called IY-Net, is designed. The proposed model has the characteristics of a fully CNN with relatively good generality. It doesn't need to specify fusion rules and has a simple network structure. This is the key innovation point. Compared with the existing CNNs and traditional fusion algorithms, the proposed model not only has generality and stability but also has some strengths in subjective visualization and objective evaluation. A deep CNN was used in [Mathiyalagan 2018] to extract the high frequency details from the two source images. A focus map is generated after the several convolution and max-pooling layers which contains the clarity information of the source images. The proposed bilateral filter is a very efficient edge-preserving filter which smoothen the regions around the boundaries of the obtained decision map. The pixel-wise weighted average strategy is calculated to get the fused image with high visual quality. Experimental results show that the proposed CNN-based method produces more natural effect of the fused image.

3.1.1.2 Fusion rules

To combine information from the obtained sub-images, various fusion rules can be used, some of them being recalled below.

Minimum/Maximum techniques: The pixels in the fused image are obtained by selecting the pixel with the lowest/highest intensity values of the pixels to form the different images [Meher 2019, James 2014].

Principal Component Analysis (PCA): PCA is a technique that converts a set of correlated variables into a set of uncorrelated variables known as principal components. The first principal component aims to explain as much of the data's variance as possible, and each succeeding component accounts for as much of the remaining variance as possible. The advantage of PCA is its ability to reduce a large set of variables to a smaller one while still containing most of the information that was available in the large set. Fusion techniques based on PCA were investigated in [Patil 2011, Zhao 2017].

Guided filtering: Guided filtering is a computationally efficient, edge-preserving, translation-variant operator based on a local linear model which avoids the drawbacks of bilateral filtering [He 2013]. It ensures that the significant features such as edges and textures are preserved in the fused image. The filter works

by computing a linear transformation of the guidance image within a local window around each pixel, thereby ensuring that the local linear model can capture the relationship between the guidance image and the source images. In [Geng 2019], experiments conducted on brain images demonstrate better results using guided filtering compared to Principal Component Analysis (PCA) and Multi-Resolution Singular Value Decomposition (MSVD) techniques. They highlight the advantages of guided filtering in terms of edge preservation and noise reduction, leading to clearer and more accurate fused images. A novel weighted average technique based on guided filtering was proposed in [Li 2013] to make full use of spatial consistency for fusion of the base and detail layers for multi-spectral, multi-focus, multimodal and multi-exposure images.

3.1.1.3 Fusion metrics

To assess the accuracy of any fusion method, it is essential to employ various fusion metrics. These metrics help in quantifying the performance and effectiveness of the fusion process. Common fusion metrics include:

Root Mean Square Error (RMSE): RMSE measures the square root of the average squared differences between the fused image and the reference image. It provides an indication of the overall deviation, with lower values representing better fusion quality:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[I_{\mathrm{f}}(i) - I_{\mathrm{r}}(i)\right]^2}, \tag{3.1}$$

where $I_{\mathrm{f}}(i)$ and $I_{\mathrm{r}}(i)$ are the pixel values of the fused and reference images at position i and n the total number of pixels.

Contrast-to-Noise Ratio (CNR): CNR evaluates the contrast of the fused image relative to the noise. It highlights the ability of the fusion method to distinguish between different regions, with higher values indicating better contrast. For two regions A and B in the image, CNR is defined as follows:

$$\mathrm{CNR} = \frac{|\mu_A - \mu_B|}{\sqrt{\sigma_A^2 + \sigma_B^2}}, \tag{3.2}$$

where μ_A, μ_B, σ_A and σ_B are the mean intensities and the standard deviations of the intensities in the regions A and B.

Peak Signal-to-Noise Ratio (PSNR): PSNR assesses the ratio between the maximum possible power of a signal and the power of corrupting noise, expressed in decibels. It evaluates the quality of the fused image, with higher values representing higher quality [Paramanandham 2018]:

$$\mathrm{PSNR} = 20\log\left[\max(I_{\mathrm{f}}, I_{\mathrm{r}})\right] - 10\log(\mathrm{RMSE}). \tag{3.3}$$

Structural Similarity Index (SSIM): The SSIM index measures the similarity between two images and is designed to complement the standard Mean Squared Error (MSE) or Peak Signal-to-Noise Ratio (PSNR) [Ma 2018]. It considers changes in structural information s, luminance l, and contrast c:

$$\mathrm{SSIM}(I_\mathrm{f}, I_\mathrm{r}) = l^\alpha(I_\mathrm{f}, I_\mathrm{r}) c^\beta(I_\mathrm{f}, I_\mathrm{r}) s^\gamma(I_\mathrm{f}, I_\mathrm{r}). \tag{3.4}$$

Given two images, luminance analysis compares the local mean intensities to assess brightness similarity, contrast analysis compares the local standard deviations to evaluate texture or variation, and the structure index measures the local Pearson correlation to capture the alignment of spatial patterns. Note that SSIM values range between -1 and 1, where 1 indicates a perfect similarity between the fused image and the reference image.

Universal Quality Index (UQI): This index is used to evaluate the fused image, determining how much salient information from the reference image is retained [Blasch 2008]. It is defined as:

$$\mathrm{UQI} = \left[\frac{\sigma_{I_r I_f}}{\sigma_{I_r}\sigma_{I_f}}\right]\left[\frac{2\mu_{I_r}\mu_{I_f}}{\mu_{I_r}^2 + \mu_{I_f}^2}\right]\left[\frac{2\sigma_{I_r}\sigma_{I_f}}{\sigma_{I_r}^2 + \sigma_{I_f}^2}\right], \tag{3.5}$$

where μ_{I_r}, μ_{I_f}, σ_{I_r} and σ_{I_f} are the means and the variances of images I_r and I_f, and $\sigma_{I_r I_f}$ is the covariance of these images.

3.1.2 Medical Image fusion

This section recalls briefly some fusion techniques used in the literature for different medical image modalities:

- Bhavana et al. [Bhavana 2015] proposed a novel fusion method using MRI and PET brain images based on the discrete wavelet transform. They test their model on three datasets - normal axial, normal coronal and Alzheimer's brain disease images, and obtain accurate results without losing any anatomical information in comparison with the existing techniques.

- Qiu et al. [Qiu 2017] proposed an image fusion model for CT and MR images. First, the images are both mapped into the non subsampled shearlet transform (NSST) domain. They use the absolute-maximum rule to merge high-frequency components and use a sparse representation-based approach to merge the low-frequency components.

- In [Yang 2019], the authors propose a novel multimodal medical image fusion method based on structural patch decomposition (SPD) and fuzzy logic technology. First, the SPD method is employed to extract two salient features for fusion discrimination. Next, two novel fusion decision maps referred to as incomplete fusion map and supplemental fusion map are constructed from

salient features. In this step, the supplemental map is constructed by two different fuzzy logic systems. The supplemental and incomplete maps are then combined to construct an initial fusion map. The final fusion map is obtained by processing the initial fusion map with a Gaussian filter. Finally, a weighted average approach is adopted to create the final fused image.

- Zhong et al. propose an unsupervised multi-stage deep learning framework called PAMRFuse for misaligned PAT and MRI image fusion. They start by a registration network to align the input pair images, then a self-attentive fusion network is used for fusion [Zhong 2024].
- In [Wang 2014], a multi-modal medical image fusion is developped using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients. Their method was validated on MRI-SPECT of normal brain aging, and MRI-PET images of a patient with Mild Alzheimer's Disease.

3.1.3 Conclusion

In the first part of this chapter, we provided a comprehensive literature review on image fusion techniques. This review included various methods and fusion rules, as well as different metrics used for evaluating image fusion quality. We presented numerous studies that address image fusion in both general contexts and specifically within the scope of medical imaging. In the following sections, we introduce the proposed fusion model for 2D MR and 2D US images. For MR images, a linear model using blurring and downsampling is employed due to its effectiveness in super-resolution. For US images, a denoising model addresses speckle noise. We propose a non-parametric transformation based on reproducing kernel theory to correlate the gray levels of these images.

3.2 2D-2D MR-US fusion using reproducing kernels

This section presents a new 2D-2D MR-US fusion algorithm for endometriosis surgery. This method combines the advantages of each modality, i.e., the good contrast and signal to noise ratio of the MR image and the good spatial resolution of the US image. The method is based on an inverse problem, performing a super-resolution of the MR image and a denoising of the US image. To model the relationships between the gray levels of the MR and US images, a non-parametric transformation built using the theory of reproducing kernel Hilbert spaces is introduced. A non-linear cost function is used to solve the fusion problem and is constructed using the observation models associated with the MR and US images, as well as the kernel transformation that relates the two images. Minimizing this cost function is challenging and is handled by a proximal alternating linearized minimization (PALM) algorithm. The proposed image fusion method is evaluated on synthetic data from an in vivo MR volume, as well as on experimental data acquired from a phantom with imaging characteristics similar to those of endometriosis. Both

qualitative and quantitative results allow the performance of the resulting fusion method to be appreciated.

3.2.1 Observation models

As explained previously, MRI and US are two medical modalities that have their specific strengths and weaknesses. The main purpose of obtaining a fused image from these two modalities is to take benefit from their strengths and mitigate their weaknesses. This section studies a fusion method for 2D MR and 2D US images. Considering the limitations of each modality, we construct two observation models to enhance the quality (in terms of resolution and contrast) of both images.

In the following, $\boldsymbol{y}_{\text{mri}} \in \mathbb{R}^M$ and $\boldsymbol{y}_{\text{us}} \in \mathbb{R}^N$ denote the observed MR and US images, which are the inputs of the proposed fusion algorithm. Both observed images are related to the ideal fused image through two models accounting for the degradations affecting the two modalities. In particular, the low spatial resolution of the MR image is classically modelled using a blur, a down-sampling operator [Gholipour 2010, Greenspan 2002, Manjón 2010] and additive white Gaussian noise. A denoising model was considered for the US image in order to mitigate the effect of speckle noise, modelled as an additive log-Rayleigh noise with localization parameter γ [Aysal 2007a, Gupta 2005]. Combining these super-resolution and denoising models leads to the following fusion model already investigated in [El Mansouri 2020]:

$$\begin{aligned} \boldsymbol{y}_{\text{mri}} &= \boldsymbol{S}\boldsymbol{C}\boldsymbol{x}_{\text{mri}} + \boldsymbol{n}_m, \\ \boldsymbol{y}_{\text{us}} &= \boldsymbol{x}_{\text{us}} + \boldsymbol{n}_u, \end{aligned} \tag{3.6}$$

where $\boldsymbol{x}_{\text{mri}} \in \mathbb{R}^N$ is the non-observable high-resolution vectorized MR image, $\boldsymbol{y}_{\text{mri}} \in \mathbb{R}^M$ is the low-resolution observed MR image, $\boldsymbol{n}_m \in \mathbb{R}^N$ is an independent identically distributed (i.i.d.) additive white Gaussian noise with variance σ_m^2, $\boldsymbol{C} \in \mathbb{R}^{N\times N}$ is a matrix with block circulant with circulant blocks modelling the blurring effect of the MRI point spread function (PSF) with circulant boundary conditions, $\boldsymbol{S} \in \mathbb{R}^{M\times N}$ (with $N = d_m^2$) is a decimation operator with a decimation factor d. On the other hand, $\boldsymbol{y}_{\text{us}} \in \mathbb{R}^N$ is the vectorized observed B-mode US image, $\boldsymbol{x}_{\text{us}} \in \mathbb{R}^N$ is the vectorized speckle noise-free US image and $n_u \in \mathbb{R}^N$ is an i.i.d. log-Rayleigh noise with localization parameter γ [1].

Since US and MR modalities are different by their nature and content, there is no simple correspondence between the gray levels of these images and we need to define a function to link $\boldsymbol{x}_{\text{us}}$ and $\boldsymbol{x}_{\text{mri}}$. [El Mansouri 2020] considered a linear function that links directly the US image to the MRI and the gradient of MRI. Motivated by [Roche 2001], a polynomial transformation was used in [El Mansouri 2020] to link

[1] The probability density function of a variable distributed according to a log-Rayleigh distribution ($z \sim \mathcal{LR}(\gamma)$), is $p(z) = \frac{(e^z)^2}{\gamma} \exp\left(-\frac{(e^z)^2}{2\gamma}\right) \mathbb{I}_{\mathbb{R}^+}(z)$, where $\mathbb{I}_{\mathbb{R}^+}$ is the indicator function on $\mathbb{R}^+$.

$\boldsymbol{x}_{\text{us}}$ and $\boldsymbol{x}_{\text{mri}}$ as:

$$\boldsymbol{x}_{\text{us}} = f\left(\boldsymbol{x}_{\text{mri}}, \nabla \boldsymbol{x}_{\text{mri}}^H \boldsymbol{u}\right), \tag{3.7}$$

where $f : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N$ is an unknown polynomial function of the image $\boldsymbol{x}_{\text{mri}}$, its gradient, and the US scan direction $\boldsymbol{u}$. Using this relationship between MR and US images, the image formation models in (3.6) can be rewritten as:

$$\begin{aligned} \boldsymbol{y}_{\text{mri}} &= \boldsymbol{SCx} + \boldsymbol{n}_m, \\ \boldsymbol{y}_{\text{us}} &= f(\boldsymbol{x}, \nabla \boldsymbol{x}_{\text{mri}}^H \boldsymbol{u}) + \boldsymbol{n}_u, \end{aligned} \tag{3.8}$$

where $\boldsymbol{x} \in \mathbb{R}^N$ is the unknown image to be estimated, containing relevant information from both MR and US data. The conditional distributions of $\boldsymbol{y}_{\text{mri}}$ and $\boldsymbol{y}_{\text{us}}$ given $\boldsymbol{x}$ can be determined using the noise distributions:

$$\begin{aligned} \boldsymbol{y}_{\text{mri}}|\boldsymbol{x} &\sim \mathcal{N}(\boldsymbol{SCx}, \sigma_m^2 \boldsymbol{I}_N), \\ \boldsymbol{y}_{\text{us}}|\boldsymbol{x} &\sim \prod_{i=1}^{N} \mathcal{LR}(\gamma). \end{aligned} \tag{3.9}$$

where $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, and $\mathcal{LR}$ is the log-Rayleigh distribution with localization parameter γ. Using Bayes rule and the independence between the noise vectors $\boldsymbol{n}_m$ and $\boldsymbol{n}_u$, the posterior distribution of $\boldsymbol{x}$ can be computed:

$$p(\boldsymbol{x}|\boldsymbol{y}_{\text{mri}}, \boldsymbol{y}_{\text{us}}) \propto p(\boldsymbol{y}_{\text{mri}}|\boldsymbol{x}) p(\boldsymbol{y}_{\text{us}}|\boldsymbol{x}) p(\boldsymbol{x}), \tag{3.10}$$

where $p(\boldsymbol{x})$ is the prior distribution of $\boldsymbol{x}$ and $\propto$ means "proportional to". Finally, the log-posterior distribution can be written as:

$$\begin{aligned} -\log p(\boldsymbol{x}|\boldsymbol{y}_{\text{mri}}, \boldsymbol{y}_{\text{us}}) = K + \underbrace{\frac{1}{2\sigma_m^2} ||\boldsymbol{y}_{\text{mri}} - \boldsymbol{SCx}||^2}_{\text{MRI fidelity term}} - \underbrace{\log[p(\boldsymbol{x})]}_{\text{regularization}} \\ + \underbrace{\sum_{i=1}^{N} \Big[\exp(y_{\text{us},i} - f_i(\boldsymbol{x}, \nabla \boldsymbol{x}^H \boldsymbol{u})) - \gamma(y_{\text{us},i} - f_i(\boldsymbol{x}, \nabla \boldsymbol{x}^H \boldsymbol{u}))\Big]}_{\text{US fidelity term}}, \end{aligned} \tag{3.11}$$

where $y_{\text{us},i}$ and $f_i(\boldsymbol{x}, \nabla \boldsymbol{x}^H \boldsymbol{u})$ are the i^{th} components of $\boldsymbol{y}_{\text{us}}$ and $f(\boldsymbol{x}, \nabla \boldsymbol{x}^H \boldsymbol{u})$ and K is a constant. Different prior distributions $p(\boldsymbol{x})$ have been considered in the literature to solve ill-posed problems. In this study, the classical total variation (TV) is used. Estimating $\boldsymbol{x}$ in the sense of the maximum a posteriori principle

using this TV regularization leads to the following minimization problem:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \underbrace{\frac{1}{2}\|\boldsymbol{y}_{\text{mri}} - \mathbf{S}\mathbf{C}\mathbf{x}\|^2}_{\text{MRI data fidelity}} + \underbrace{\tau_1\|\nabla\mathbf{x}\|^2 + \tau_3\|\nabla f(\boldsymbol{x}, \nabla\boldsymbol{x}_m^H\boldsymbol{u})\|^2}_{\text{regularization}}$$
$$\underbrace{+\tau_2 \sum_{i=1}^{N} \left\{\exp\left[y_{\text{us},i} - \nabla f_i(\boldsymbol{x}, \nabla\boldsymbol{x}_m^H\boldsymbol{u})\right] - \gamma\left[y_{\text{us},i} - \nabla f_i(\boldsymbol{x}, \nabla\boldsymbol{x}_m^H\boldsymbol{u})\right]\right\}}_{\text{US data fidelity}},$$

that can be solved using the proximal alternating linearized minimization (PALM) [Bolte 2014, El Mansouri 2020].

3.2.2 Relation between MR and US images

This section defines a new function f to link US and MR images based on reproducing kernels. Kernel methods have become universal since they are capable of approaching complex non linear relationships between signal and images. These methods have been used successfully in the context of point cloud registration [Ma 2016], image denoising [Bouboulis 2010], super-resolution [Deng 2016] and segmentation [Kang 2014]. They need to define an appropriate kernel function between the signals or images of interest such as the Gaussian kernel with parameter β defined as:

$$K(\boldsymbol{x}, \boldsymbol{y}) = \exp\left(-\beta\|\boldsymbol{x} - \boldsymbol{y}\|^2\right), \tag{3.12}$$

where $\|.\|$ is an appropriate norm. In the following, as in [Bouboulis 2010], $\boldsymbol{x}$ and $\boldsymbol{y}$ are patches of size $n \times n$ centered around pixels of the US and MR images. The idea is that each pixel of a patch in the noise-free US image is a linear combination of the kernel applied to the pixels of the corresponding MR image patch. Since a given pixel of the US image belongs to several patches, the final value of the pixel intensity of the US image is defined as the average of the intensities resulting from different patches, i.e.,

$$\boldsymbol{x}_{\text{us},i} = \frac{1}{n_i} \sum_{p \in J_i} \left[\sum_{k=1}^{n^2} c_{p,k} K(\boldsymbol{x}_{\text{mri},i}, \boldsymbol{x}_{\text{mri},h_k(p)})\right], \tag{3.13}$$

where $\boldsymbol{x}_{\text{us},i}$ is the ith US pixel (belonging to different patches of size n^2), J_i is the set of indices of the patches containing the pixel i, $n_i \leq n^2$ is the cardinality of J_i (each patch is identified by the index of its central pixel) and $h_k(p)$ is the index of the kth pixel of Patch $\#p$. To estimate the coefficient vector $\boldsymbol{c}_p = (c_{p,1}, ..., c_{p,n^2})^T$ of Patch $\#p$, we assume that the transformation (3.13) is also valid for the observed MR and US images and consider the least squares (LS) estimator

$$\min_{\boldsymbol{c}_p} \sum_{i=1}^{n^2} \left[\boldsymbol{y}_{\text{us},i} - \sum_{k=1}^{n^2} c_{p,k} K(\boldsymbol{y}_{\text{mri},i}, \boldsymbol{y}_{\text{mri},h_k(p)})\right]^2, \tag{3.14}$$

that can be written in matrix form as:

$$\min_{\boldsymbol{c}_p} ||\boldsymbol{A}\boldsymbol{c}_p - \boldsymbol{p}_{\text{us}}||^2, \tag{3.15}$$

where $A_{i,j} = K(\boldsymbol{p}_{\text{mri}_i}, \boldsymbol{p}_{\text{mri}_j})$, $\boldsymbol{p}_{\text{us}}$ and $\boldsymbol{p}_{\text{mri}}$ are patches extracted from $\boldsymbol{y}_{\text{us}}$ and $\boldsymbol{y}_{\text{mri}}$. The solution of (3.15) is $\hat{\boldsymbol{c}}_p = \boldsymbol{A}^{\dagger}\boldsymbol{p}_{\text{us}}$, where $\boldsymbol{A}^{\dagger} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T$ is the pseudo-inverse of $\boldsymbol{A}$. Another strategy would be to plug the relation (3.13) into the observation models (3.6) and to estimate jointly the unknown image $\boldsymbol{x}_{\text{mri}}$ and the coefficient vector $\boldsymbol{c}_p$ using a modified PALM algorithm.

3.2.3 PALM Algorithm

The presence of the nonlinear function f in the minimization problem prevents the use of algorithms based on the alternate direction method of multipliers (ADMM). Consequently, we propose hereafter an algorithm based on the proximal alternating linearized minimization (PALM), adapted to nonconvex and nonsmooth functions.

The PALM algorithm was originally designed to minimize functions of two vectors $\boldsymbol{x}$ and $\boldsymbol{v}$ for the following optimization problem:

$$\min_{\boldsymbol{x},\boldsymbol{v}} \psi(\boldsymbol{x}, \boldsymbol{v}) := l(\boldsymbol{x}) + g(\boldsymbol{v}) + H(\boldsymbol{x}, \boldsymbol{v}), \tag{3.16}$$

where l and g are continuous convex functions and H may be non-linear. Moreover, these three functions must respect the following conditions to fit the PALM framework:

- l and g are inf-bounded $\inf_{\mathbb{R}^N}(l) > -\infty$ et $\inf_{\mathbb{R}^N}(g) > -\infty$
- For any fixed $\boldsymbol{v}$, the function $\boldsymbol{x} \mapsto H(\boldsymbol{x}, \boldsymbol{v})$ is $C^{1,1}$ and the partial gradient $\nabla_{\boldsymbol{x}} H(\boldsymbol{x}, \boldsymbol{v})$ is globally Lipschitz.
- For any fixed $\boldsymbol{x}$, the function $\boldsymbol{v} \mapsto H(\boldsymbol{x}, \boldsymbol{v})$ is $C^{1,1}$ and the partial gradient $\nabla_{\boldsymbol{v}} H(\boldsymbol{x}, \boldsymbol{v})$ is globally Lipschitz.
- ∇H is Lipschitz continuous on bounded subsets of the image domain.

PALM can be viewed as a minimization of the sum of the two functions l and g with a linearization of H around a given point $\boldsymbol{x}^k$. The alternate minimization with respect to the two blocks $\boldsymbol{x}$ and $\boldsymbol{v}$ proposed in [Bolte 2014] generates a sequence $(\boldsymbol{x}^k, \boldsymbol{x}^k)$ using the following steps:

Step 1: Choose $\gamma_1 > 1$, set $c_k = \gamma_1 L_{\boldsymbol{x}}(\boldsymbol{v}^k)$ and update $\boldsymbol{x}^k$ as follows:

$$\begin{aligned}\boldsymbol{x}^{k+1} &= \text{prox}^{l}_{c_k}\left(\boldsymbol{x}^k - \frac{1}{c_k}\nabla_{\boldsymbol{x}} H(\boldsymbol{x}^k, \boldsymbol{v}^k)\right), \\ &= \text{argmin}_{\boldsymbol{x}}\ (\boldsymbol{x} - \boldsymbol{x}^k)^H \nabla_{\boldsymbol{x}} H(\boldsymbol{x}^k, \boldsymbol{v}^k) + \frac{c_k}{2}||\boldsymbol{x} - \boldsymbol{x}^k||^2 + l(\boldsymbol{x}),\end{aligned}$$

where $L_{\boldsymbol{x}}(\boldsymbol{v}^k)$ is the Lipschitz constant of $\boldsymbol{x} \mapsto H(\boldsymbol{x}, \boldsymbol{v}^k)$ and $\boldsymbol{v}^k$ is the value of $\boldsymbol{v}$ at iteration k.

Step 2: Choose $\gamma_2 > 1$, set $d_k = \gamma_2 L_{\boldsymbol{v}}(\boldsymbol{x}^k)$ and update $\boldsymbol{v}^k$ as follows:

$$\begin{aligned}\boldsymbol{v}^{k+1} &= \text{prox}^g_{d_k}\left(\boldsymbol{v}^k - \frac{1}{d_k}\nabla_{\boldsymbol{v}} H(\boldsymbol{x}^k, \boldsymbol{v}^k)\right), \\ &= \text{argmin}_{\boldsymbol{v}}\ (\boldsymbol{v} - \boldsymbol{v}^k)^H \nabla_{\boldsymbol{v}} H(\boldsymbol{x}^k, \boldsymbol{v}^k) + \frac{d_k}{2}||\boldsymbol{v} - \boldsymbol{v}^k||^2 + g(\boldsymbol{x}),\end{aligned}$$

where $L_{\boldsymbol{v}}(\boldsymbol{x}^k)$ is the Lipschitz constant of $\boldsymbol{v} \mapsto H(\boldsymbol{x}^k, \boldsymbol{v})$ and $\boldsymbol{x}^k$ is the value of $\boldsymbol{x}$ at iteration k.

3.2.4 PALM for MR-US fusion

This section studies a PALM algorithm allowing (3.6) to be solved, when the US and MR images are linked by (3.13). We introduce the following functions:

$$l(\boldsymbol{x}) = \frac{1}{2}||\boldsymbol{y}_{\text{mri}} - \boldsymbol{SC}\boldsymbol{x}||^2 + \tau_1||\nabla \boldsymbol{x}||^2, \tag{3.17}$$

$$g(\boldsymbol{v}) = \tau_2 \sum_{i=1}^{N} \left[\exp(y_{\text{us},i} - v_i) - \gamma(y_{\text{us},i} - v_i)\right] + \tau_3||\nabla \boldsymbol{v}||^2, \tag{3.18}$$

$$H(\boldsymbol{x}, \boldsymbol{v}) = \tau_4 \sum_{i=1}^{N} \left(v_i - \frac{1}{n_i} \sum_{p \in J_i} \left[\sum_{k=1}^{n^2} c_{p,k} K(x_i, x_{h_k(p)}) \right] \right)^2 \tag{3.19}$$

where $\boldsymbol{v} = f(\boldsymbol{x})$. Note that l and g are data fidelity terms associated with the MR and US images and that H is used to enforce the relationship between the US and MR images based on kernels (defined in (3.13)). Using these definitions, the PALM algorithm reduces to alternate between updates of $\boldsymbol{x}$ and $\boldsymbol{v}$ as explained below.

3.2.4.1 Update $\boldsymbol{x}$

The first step of PALM algorithm reduces to

$$\begin{aligned}\boldsymbol{x}^{k+1} &= \text{prox}^l_{L_{k+1}}\left(\boldsymbol{x}^k - \frac{1}{L_{k+1}}\nabla_{\boldsymbol{x}} H(\boldsymbol{x}^k, \boldsymbol{v}^k)\right), \\ &= \text{argmin}_{\boldsymbol{x}} \left\{ \frac{1}{2}||\boldsymbol{SC}\boldsymbol{x} - \boldsymbol{y}_{\text{mri}}||^2 + \tau_1||\nabla \boldsymbol{x}||^2 \right. \\ &\quad \left. + \frac{L_{k+1}}{2}\left\|\boldsymbol{x} - \left(\boldsymbol{x}^k - \frac{1}{L_{k+1}}\nabla_{\boldsymbol{x}} H(\boldsymbol{x}^k, \boldsymbol{v}^k)\right)\right\|^2 \right\},\end{aligned}$$

with L_{k+1} the Lipschitz constant at iteration $k+1$. This minimization problem admits an analytical solution, which can be computed efficiently in the Fourier

domain. The update of $\boldsymbol{x}$ at iteration $k+1$ is obtained as follows:

$$\boldsymbol{x}^{k+1} = \left[\boldsymbol{C}^H\boldsymbol{S}^H\boldsymbol{S}\boldsymbol{C} + 2\left(\tau_1\boldsymbol{D} + \frac{L_{k+1}}{2}\boldsymbol{I}_N\right)\right]^{-1}\boldsymbol{R}, \tag{3.20}$$

with

$$\boldsymbol{D} = \boldsymbol{D}_h^H\boldsymbol{D}_h + \boldsymbol{D}_v^H\boldsymbol{D}_v$$

and

$$\boldsymbol{R} = \boldsymbol{C}^H\boldsymbol{S}^H\boldsymbol{y}_{\text{mri}} + 2L_{k+1}[\boldsymbol{x}^k - \frac{1}{L_{k+1}}\nabla_{\boldsymbol{x}}H(\boldsymbol{x}^k, \boldsymbol{v}^k)],$$

where $\boldsymbol{D}_h$ and $\boldsymbol{D}_v$ are the horizontal and vertical finite difference operators, $\boldsymbol{n} = (n_1, ..., n_N)^T$ and $\boldsymbol{J} = (J_1, ..., J_N)^T$ gathers all the patches. The direct computation of (3.20) requires the inversion of a high-dimensional matrix, which can be handled using a diagonalization in the Fourier domain [El Mansouri 2020].

Appendix A further details the computation of the Lipschitz constant and the update of $\boldsymbol{x}$.

3.2.4.2 Update $\boldsymbol{v}$

The vector $\boldsymbol{v}$ is updated using gradient descent with backtracking line search, given that the function to be minimized here is differentiable and convex, i.e.,

$$\begin{aligned}\boldsymbol{v}^{k+1} = \text{argmin}_{\boldsymbol{v}} \quad & \tau_2\sum_i[\exp(y_{\text{us},i} - v_i) - \gamma(y_{\text{us},i} - v_i)] + \\ & \tau_3||\nabla\boldsymbol{v}||^2 + \frac{d_k}{2}\left\|\boldsymbol{v} - \left(\boldsymbol{v}^k - \frac{1}{d_k}\nabla_{\boldsymbol{v}}H(\boldsymbol{x}^{k+1}, \boldsymbol{v}^k)\right)\right\|^2,\end{aligned} \tag{3.21}$$

with $d_k = L_{\boldsymbol{v}}(\boldsymbol{x}^k) = 2\tau_4$ the Lipschitz constant of $\nabla_{\boldsymbol{v}}H$.

3.2.4.3 Proposed fusion algorithm

Algorithm 1 summarizes the different steps of the proposed fusion algorithm using reproducing kernels:

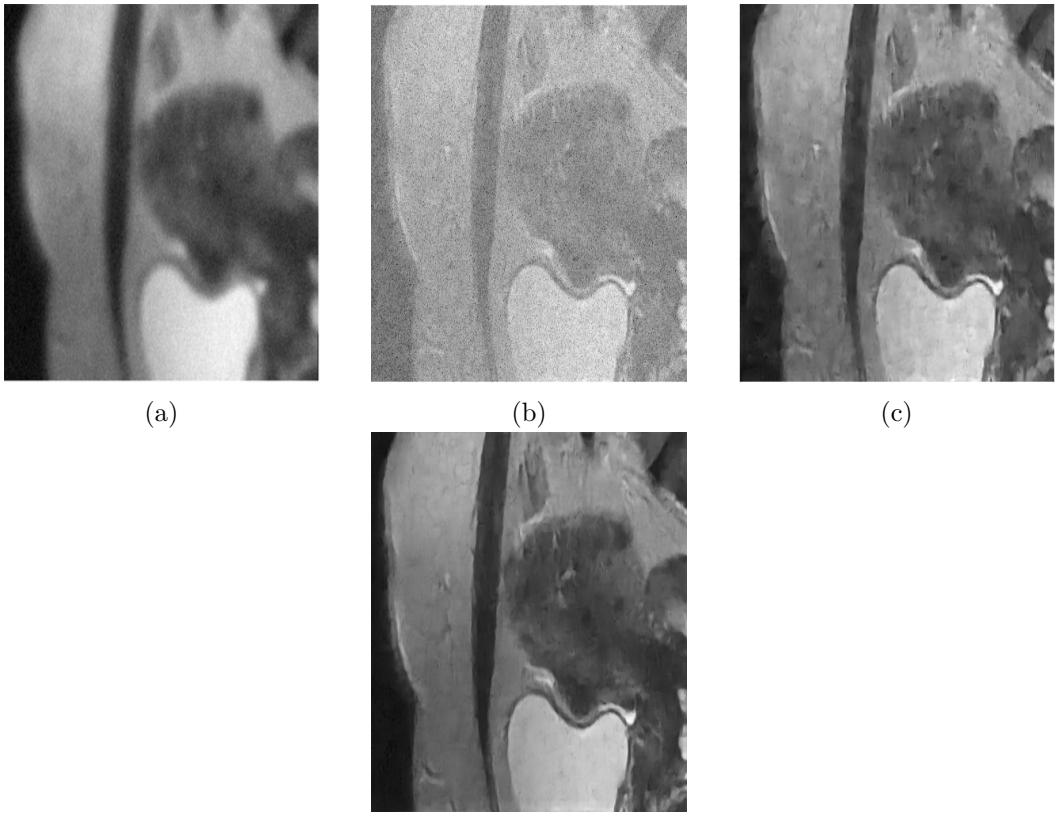
Algorithm 1 Proposed Kernel-based Fusion Algorithm.

1: Input: $\boldsymbol{y}_m$, $\boldsymbol{y}_u$, $\boldsymbol{S}$, $\boldsymbol{C}$, τ, γ, β
2: Estimate the coefficients of the kernel as $\hat{\boldsymbol{c}}_p = \boldsymbol{A}^\dagger\boldsymbol{p}_{\text{us}}$
3: **while** stopping criterion is not satisfied **do**
4: **Step 1 :** Compute L_{k+1} and update $\boldsymbol{x}$ using (3.20)
5: **Step 2 :** Set $d_k = 2\tau_4$ and update $\boldsymbol{v}$ using (3.21)
6: **end while**
7: Output: Fused image $\boldsymbol{x}$

3.3 Experiments

3.3.1 Synthetic data from real MR acquisition

The proposed MRI/US fusion algorithm was first validated on synthetic data. The simulations presented hereafter have been obtained using a real high resolution MR image that has been degraded to generate an image close to that obtained for endometriosis surgery. The 3D high resolution MR volume corresponds to a real pelvic MRI capturing the uterus, bladder and endometriosis lesions. A blurred and noisy 3D MRI is then generated from this high-resolution MR volume. More precisely, the HR volume was blurred using a 2D Gaussian filter of standard deviation $\sigma^2 = 4$, and then was contaminated by an additive white Gaussian noise (SNR = 21.5 dB), yielding the MR image displayed in Fig. 3.4(a). A kernel transformation as defined in (3.13) was used to generate the clean US image from the corresponding clean high-resolution MR image. Then, log-Rayleigh additive noise was added, yielding the image displayed in Fig. 3.4(b) (SNR= 11.5 dB).



(a) (b) (c)

(d)

Figure 3.4: (a) MR image, (b) US image, (c) MARIUS image: Proposed approach, (d) MARIUS image: Polynomial approach [El Mansouri 2020].

The performance of the non-parametric fusion method was evaluated qualitatively through visual inspection of the final image, and quantitatively using the

contrast-to-noise ratio (CNR). For two patches extracted from two different structures (uterus and bladder in this case), we recall that the CNR is defined as $\text{CNR} = \frac{|\mu_i - \mu_j|}{\sqrt{\sigma_i^2 + \sigma_j^2}}$, where $\mu_i, \mu_j, \sigma_i^2, \sigma_j^2$ are the means and standard deviations of two blocks of pixels.

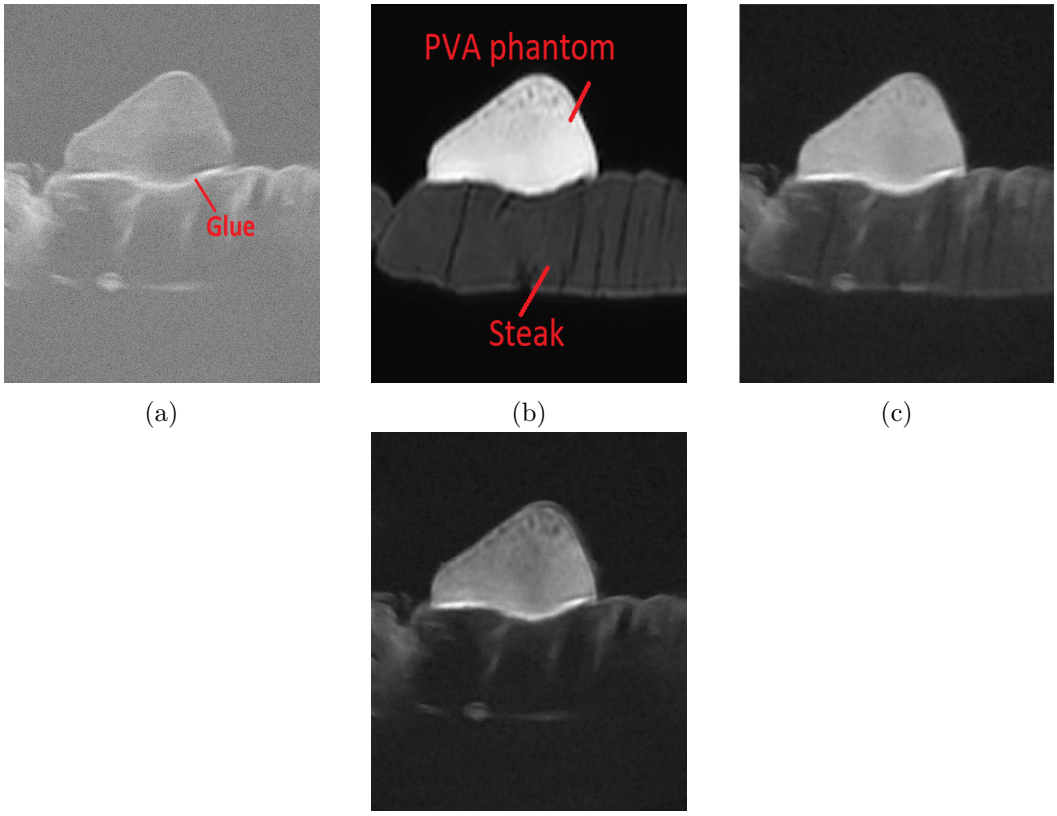
The MARIUS image obtained using the proposed algorithm is displayed in Fig. 3.4(c). This image provides a good compromise between the US and MR data. Specifically, the fused image is less affected by US speckle and MRI blur, provides well-defined contours and good contrast compared to the native MR and US images. The CNR values for the MR and the US images are 39.52 dB and 27.15 dB, whereas the MARIUS images obtained using polynomial (Fig. 3.4(d)) and kernel transformations have CNRs equal to 34.96 dB and 38.65 dB respectively. These results show the interest of replacing the polynomial by a kernel.

3.3.2 Phantom data

This section evaluates the proposed fusion algorithm on experimental phantom data. The phantom was made of a beaf steak on top of which was stuck a polyvinyl alcohol (PVA) phantom using cyanoacrylate instant glue. It was designed to mimic uterus and endometrium responses to MR and US imaging. More details about the experimental model design and image acquisition can be found in [Vidal 2019]. The sizes of the acquired images are (600×600) for the US, and $(320 \times 320 \times 90)$ for the MRI volume. The field of view of the MR image is wider than the one of the US image. Therefore, the MR volume was manually cropped to $(100 \times 100 \times 90)$ to ensure similar fields of view for the two modalities. Bicubic interpolation of the MR image was finally performed to ensure the same pixel size in MR and US images. Note that the proposed fusion method requires registered images. In the following experiments, the images were registered using the algorithm presented in Section 2.2 and are shown in Figs. 3.5(a,b).

The proposed fusion algorithm was applied to the US and MR images with the following parameters: $\beta = 10^{-5}$, $\tau_1 = 10^{-5}, \tau_2 = 0.5, \tau_3 = 0.01$ and $\tau_4 = 10^{-5}$. The parameters were fixed to their best values by visual inspection of the fused images. Section 3.3.2.2 details the influence of the hyperparameters on the fused image.

The first result is that the fused image shown in Fig. 3.5(c) has a good spatial resolution similar to the US image, and a contrast equivalent to the MRI. In particular, the different structures of interest are much better highlighted: (i) the glue between the steak and the PVA phantom, mimicking the depth of penetration, not visible on MRI because of the lack of resolution, appears clearly on the US and MARIUS images, this part is really crucial for the case of endometriosis surgery since it could represent the border between a tumor and the uterine tissue in which it is necessary to incise with precision. Making this part of the image stand out confirms the efficiency of the model; (ii) the steak and the PVA are well contrasted on the MR and MARIUS images, which is not the case in the US image.



(a) (b) (c)

(d)

Figure 3.5: (a) US image, (b) MR image, (c) MARIUS image: Proposed approach, (d) MARIUS image: Polynomial approach [El Mansouri 2020].

This improved resolution is visually depicted in Fig. 3.6, which shows pixel intensities for the same vertical straight line (Fig. 3.7) from the US, MR and MARIUS images using a kernel transformation (proposed method) and a polynomial [El Mansouri 2020]. For the MRI, the plot shows the change of pixel intensities between the beef steak and the PVA, this indicates the presence of contrast but a less defined edge. For the US, one can see the sharp transition in pixel intensities at the boundary, highlighting US'ability to detect fine details and precise edges. In the fused image, both the intensity difference and the peak are present, combining the strenghts of both modalities. Quantitative results are provided in Tab. 4.1, which reports the CNR values between the PVA phantom and the beef steak for the different images.

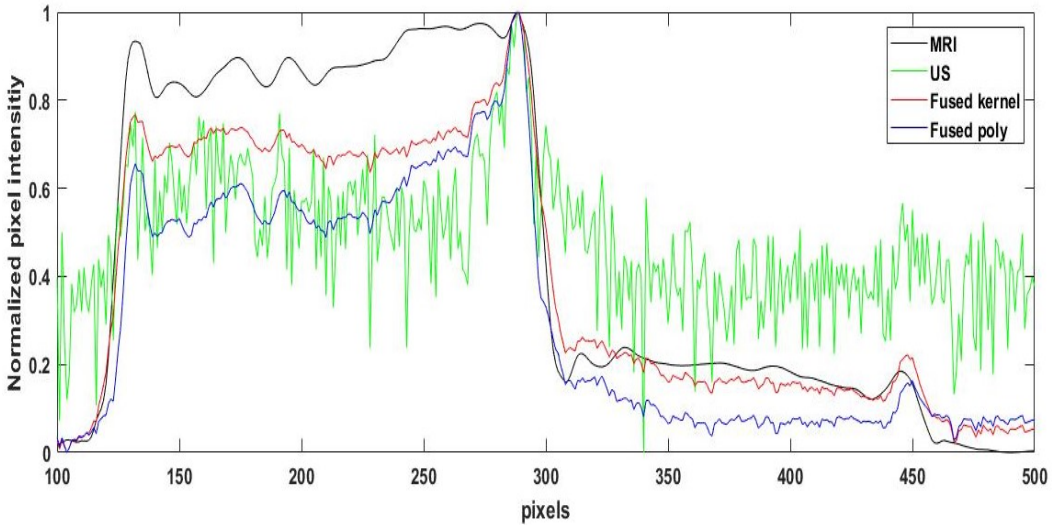


Figure 3.6: Normalized pixel intensities of US (Green), MRI (Black), and MARIUS images using polynomial (Blue) and kernel functions (Red).

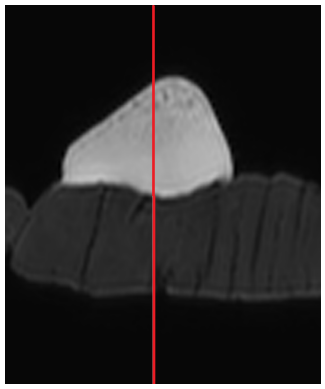


Figure 3.7: The vertical straight line.

	MRI	US	Proposed	Fusion [El Mansouri 2020]
CNR	54.21 dB	18.91 dB	52.23 dB	43.17 dB

Table 3.1: CNR values for the US, MR and MARIUS images.

3.3.2.1 Influence of β

This section analyzes the impact of the parameter β of the Gaussian kernel on the fusion results. The value of β is related to the ability of the kernel approximation to reconstruct the US image from the MR image. In order to appreciate the role of this parameter on the fusion results, problem (3.14) was solved for different values of β. The root mean square error (RMSE) between the true US image and

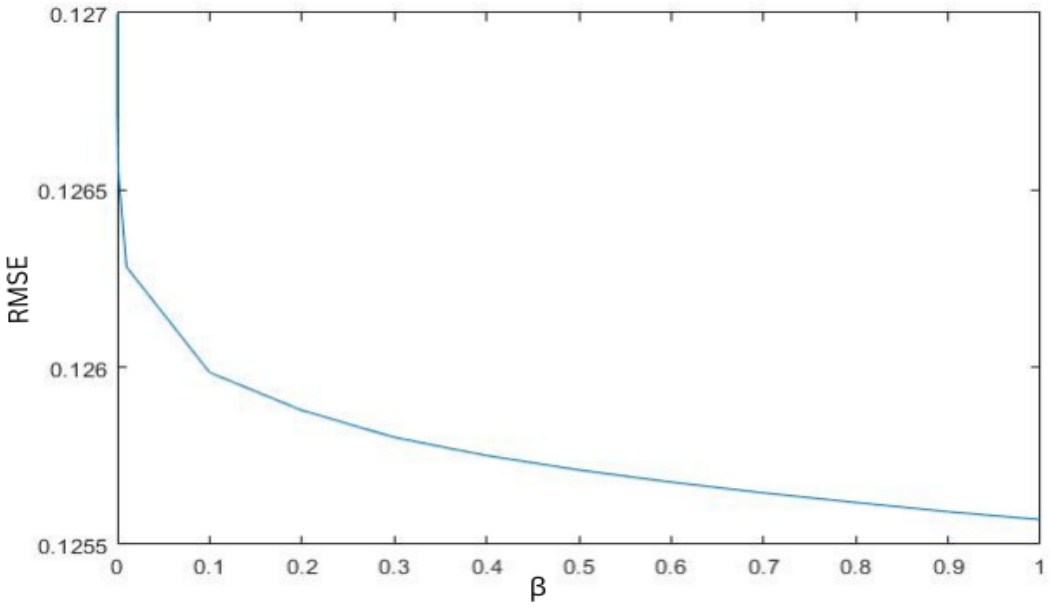


Figure 3.8: RMSE between the image estimated using a non-parametric kernel transformation and the targeted US image.

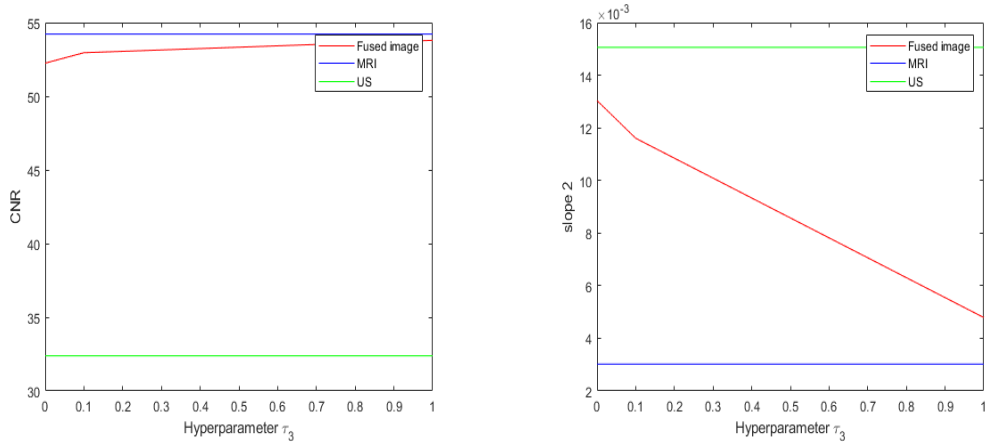
its reconstruction obtained using the MRI as defined in 3.13 was computed as a function of β yielding Fig. 3.8. One may note from Fig. 3.8 that the RMSE is a decreasing function of β, which is explained in this section.
First, note that when β tends to infinity, one has:

$$\forall i \neq j,\ K(\boldsymbol{p}_{\text{mri}_i}, \boldsymbol{p}_{\text{mri}_j}) = \exp(-\beta(\boldsymbol{p}_{\text{mri}_i}, \boldsymbol{p}_{\text{mri}_j})) \underset{\beta \to +\infty}{\longrightarrow} 0.$$

and thus the matrix $\boldsymbol{A}$ tend towards the identity when β tends to infinity. Therefore, for large values of β, the problem $\min_{c_p} ||Ac_p - \boldsymbol{p}_{\text{us}}||^2$ becomes trivial and admits the solution $c_p = \boldsymbol{p}_{\text{us}}$, implying $\boldsymbol{p}_{\text{us, estimated}} = Ac_p = \boldsymbol{p}_{\text{us}}$. An undesirable effect occurs in this case, by perfectly modeling the relationship between MR and US pixels. At the other extreme, when β tends towards 0, all the elements in A tend to 1: $A^\dagger = \frac{1}{n^4} A$, so $c_p = A^\dagger \boldsymbol{p}_{\text{us}} = \frac{1}{n^4} A \boldsymbol{p}_{\text{us}} = \frac{1}{n^4} \sum_{i=1}^{n^2} \boldsymbol{p}_{\text{us},i} \, (1, ..., 1)^T$ and $\boldsymbol{p}_{\text{us, estimated}} = Ac_p = \frac{1}{n^2} \sum_{i=1}^{n^2} \boldsymbol{p}_{\text{us},i} \, (1, ..., 1)^T$. Each estimated patch is in this case constant, equal to the average intensity of the target US patch. Note that the fusion algorithm fails when β is beyond 10^{-2} due to a gradient calculation in (3.21) that becomes NaN. In conclusion, the best value of β should be neither too large nor too small. In this study, β was set to 10^{-5}.

3.3.2.2 Other hyperparameters

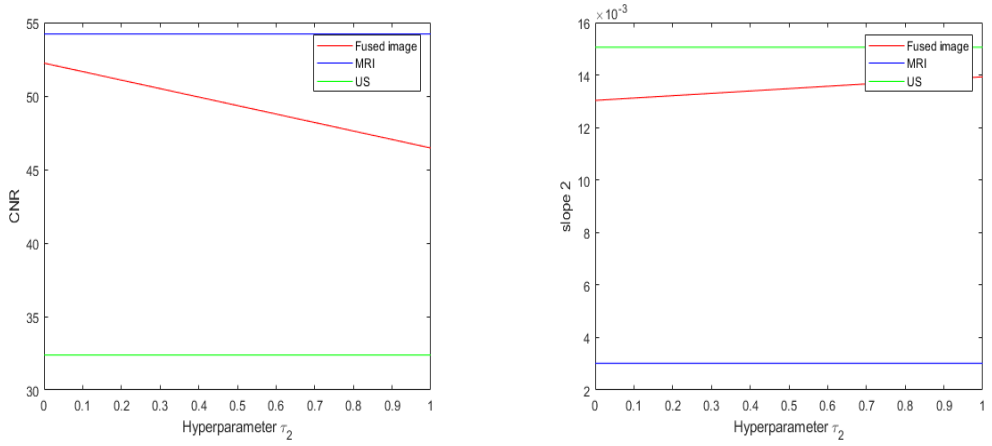
This section shows the effect of the different hyperparameters τ_1, τ_2, τ_3 and τ_4 of the proposed algorithm on fusion results. The hyperparameters τ_1 and τ_3 have an influence on the resolution of the image, which means that increasing τ_1 or τ_3



(a) (b)

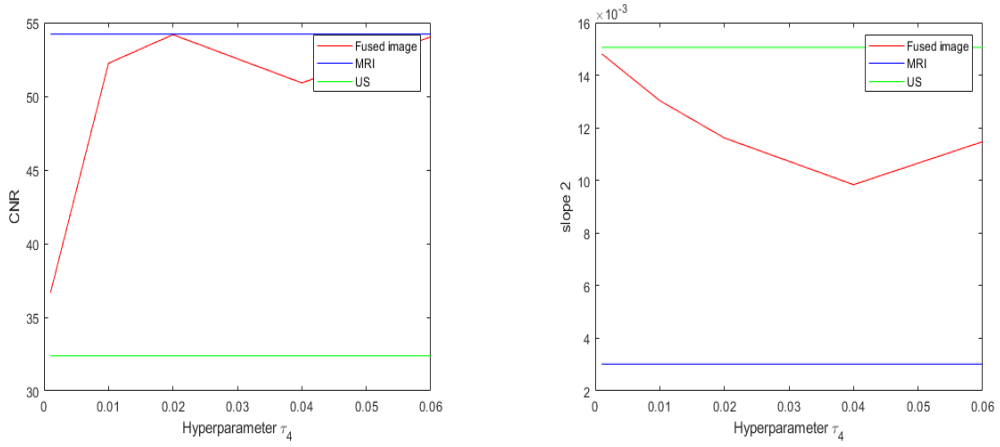
Figure 3.9: Influence of the parameter τ_3 on the fused image. (a) shows the CNR evolution, (b) shows the slope evolution.

decreases the resolution of the fused image, that is measured using slope interfaces. Fig. 3.9 of the evolution of the parameter τ_3 highlights this: When τ_3 increases the slope starts decreasing and the fused image is blurred.
τ_2 and τ_4 weight the dependence of the fused image on the MR or US image. Fig. 3.10 and 3.11 show the influence of the parameters τ_2 and τ_4 on the CNR and slope values for the US, MR and fused images. Increasing (or decreasing) these hyperparameters make the fused image close to the MRI or (US image).



(a) (b)

Figure 3.10: Influence of the parameter τ_2 on the fused image. (a) shows the CNR evolution, (b) shows the slope evolution.



(a) (b)

Figure 3.11: Influence of the parameter τ_4 on the fused image. (a) shows the CNR evolution, (b) shows the slope evolution.

3.4 Conclusion

This chapter studied a new fusion method based on reproducing kernels for MR and US images. The fused MARIUS image obtained with this method brings together the advantages of both modalities: resolution and contrast. The MARIUS image is more contrasted when using a kernel transformation instead of a polynomial. This improved contrast is clearly interesting for the detection and treatment of lesions related to endometriosis. Another significant advantage in favour of the kernel transformation is that it is not directly related to the direction of propagation of the scan US (vector $\boldsymbol{u}$ in [El Mansouri 2020]), which is not easy to obtain in practical applications. The price to pay with the proposed approach is its computational complexity. Indeed, unlike the polynomial approach which requires only about ten coefficients to be estimated upstream, the kernel transformation may need some hundreds of thousands of parameters to be estimated depending on the image and patch sizes.

In the next chapter, we propose a novel fusion method that emphasizes a two-scale decomposition to highlight the advantages of each imaging modality. This approach aims to extract and enhance the strengths of both images. Additionally, we compute specific weights to apply to these decomposed images, creating a final fused image that is more informative and comprehensive.

Chapter 4

2D/2D MR-US fusion using Guided filtering

Contents

This chapter studies a new fusion method designed for magnetic resonance (MR) and ultrasound (US) images. The proposed method is based on guided filtering, leveraging the advantages of this technique to enhance the quality of fused images. The fused image is a weighted average of base and detail images computed from the MR and US images. The weights assigned to the US image account for the presence of speckle noise, a common challenge in US imaging whereas the weights assigned to the MR image allow the contrast of the fused image to be enhanced. The effectiveness of the method is evaluated using synthetic, phantom and real data, showing promising results. The image provided by the proposed fusion method holds potential for enhancing visualization and aiding decision-making in endometriosis surgery, offering a valuable contribution to the field of medical image fusion.

4.1 Image Guided Filtering

Using the notations of [Li 2013], the guided filter involves a guidance image $\boldsymbol{I} = (I_i)$ and an input image $\boldsymbol{P} = (P_i)$ and provides an output image $\boldsymbol{O} = (O_i)$. All these images are vectorized and i denotes the pixel index. The guided filter assumes that

the output $\boldsymbol{O}$ is a linear transformation of the guidance image $\boldsymbol{I}$ in a local window ω_k, i.e.,

$$O_i = a_k I_i + b_k, \qquad \forall i \in \omega_k, \tag{4.1}$$

where ω_k is a window of radius r centered around the pixel k, and (a_k, b_k) are some coefficients assumed to be constant in ω_k, that can be estimated by minimizing the squared difference between the output image $\boldsymbol{O}$ and the input image $\boldsymbol{P}$:

$$E(a_k, b_k) = \sum_{i \in \omega_k} \left[(a_k I_i + b_k - P_i)^2 + \epsilon a_k^2 \right], \tag{4.2}$$

where ϵ is a regularization parameter that needs to be adjusted by the user, preventing a_k from being too large. Straightforward computations detailed in [He 2013] show that the coefficients a_k and b_k have the following closed-form expressions:

$$\begin{aligned} a_k &= \frac{\frac{1}{|\omega_k|} \sum_{i \in \omega_k} I_i P_i - \mu_k \bar{P}_k}{\delta_k + \epsilon}, \\ b_k &= \bar{P}_k - a_k \mu_k, \end{aligned} \tag{4.3}$$

where μ_k and δ_k are the mean and variance of the guidance image $\boldsymbol{I}$ in ω_k, $|\omega_k|$ is the number of pixels in ω_k, $\bar{P}_k$ is the mean of the input image $\boldsymbol{P}$ in ω_k, and ϵ represents the blur degree of the filter. Considering that pixel i is involved in all the overlapping windows ω_k that cover i, the value of O_i in (4.1) is not identical when it is computed in different windows. A simple strategy is to average all the possible values of O_i, which means that, after computing (a_k, b_k) for all windows ω_k in the image, the filtered output is:

$$O_i = \frac{1}{|\omega_k|} \sum_{k | i \in \omega_k} (a_k I_i + b_k). \tag{4.4}$$

Finally, according to [He 2013], (4.4) can be further converted as

$$O_i = \sum_j W_{i,j}(I_i) P_j, \tag{4.5}$$

where $W_{i,j}$ is a function of the guidance image $\boldsymbol{I}$, assumed independent of $\boldsymbol{P}$. The guided filter operation will be denoted as $\boldsymbol{O} = \mathrm{GF}_{r,\epsilon}(\boldsymbol{P}, \boldsymbol{I})$.

4.2 Image Fusion with a Guided Filter

The image fusion model introduced in this section assumes that the MR and US images to be fused, denoted as $\boldsymbol{I}_{\mathrm{mri}}$ and $\boldsymbol{I}_{\mathrm{us}}$, are already aligned, i.e., there is no geometric distortion between them, $\boldsymbol{I}_{\mathrm{mri}}$ is a 2D image extracted from the MRI that best matches $\boldsymbol{I}_{\mathrm{us}}$. This can be obtained after an appropriate pre-registration [Bennioui 2023]. Inspired from prior work in image fusion using guided filtering [Li 2013], the proposed algorithm is based on MR and US image decompositions

and subsequent weight computations. First, two-scale representations are obtained by applying an average filter to the MR and US images yielding base and detail images. Then, the base and detail MR and US images are weighted. A novel aspect is introduced by incorporating speckle-related considerations in the weight assignment for the US image.

4.2.1 Proposed fusion algorithm

Algo. 2 summarizes the different steps of the fusion model.

Algorithm 2 Fusion of MR and US images.

1: Input: $\boldsymbol{I}_{\text{us}}$, $\boldsymbol{I}_{\text{mri}}$ and parameters r_1, ϵ_1, r_2, ϵ_2, patch size $|\omega|$, Z, K_1 and K_2
2: Output: Fused Image $\boldsymbol{F}$
3: Step 1: Two-scale decompositions of $\boldsymbol{I}_{\text{us}}$, $\boldsymbol{I}_{\text{mri}}$ with (4.7)
4: Step 2: Calculate $\bar{\boldsymbol{B}}_{\text{mri}}$ and $\bar{\boldsymbol{D}}_{\text{mri}}$ as explained in (4.9)
5: Step 3: Calculate $\bar{\boldsymbol{B}}_{\text{us}}$ and $\bar{\boldsymbol{D}}_{\text{us}}$:
   a) Construct guidance images $\boldsymbol{B}^g_{\text{us}}$ and $\boldsymbol{D}^g_{\text{us}}$ using 3×3 averaging of $\boldsymbol{B}_{\text{us}}$ and $\boldsymbol{D}_{\text{us}}$.
   b) Compute $W_{i,j}(\boldsymbol{B}_{\text{us}}, \boldsymbol{B}^g_{\text{us}})$, $W_{i,j}(\boldsymbol{D}_{\text{us}}, \boldsymbol{D}^g_{\text{us}})$ in (4.12)
   c) Compute $\bar{\boldsymbol{B}}_{\text{us}}$ and $\bar{\boldsymbol{D}}_{\text{us}}$ in (4.13)
6: Step 4: Combine the base and detail layers for fusion

$$\boldsymbol{F} = \bar{\boldsymbol{B}}_{\text{mri}} + \bar{\boldsymbol{D}}_{\text{mri}} + \bar{\boldsymbol{B}}_{\text{us}} + \bar{\boldsymbol{D}}_{\text{us}}. \tag{4.6}$$

4.2.2 Two-scale decomposition

The base and detail layers of each image are obtained as:

$$\begin{aligned} \boldsymbol{B}_{\text{mri}} &= \boldsymbol{I}_{\text{mri}} * Z \quad ; \quad \boldsymbol{B}_{\text{us}} = \boldsymbol{I}_{\text{us}} * Z \\ \boldsymbol{D}_{\text{mri}} &= \boldsymbol{I}_{\text{mri}} - \boldsymbol{B}_{\text{mri}} \quad ; \quad \boldsymbol{D}_{\text{us}} = \boldsymbol{I}_{\text{us}} - \boldsymbol{B}_{\text{us}} \end{aligned} \tag{4.7}$$

where Z is an average filter [Gonzalez 2008] and $*$ denotes 2D convolution. This two-scale decomposition is designed to partition each source image into a base layer capturing the significant variations in intensity at a larger scale, and a detail layer containing the finer details at a smaller scale.

4.2.3 Weight construction for the MR image

The weights for the MR are constructed as in [Li 2013]. First, a Laplacian filter is applied to each source image to obtain the high-pass filtered images:

$$\boldsymbol{H}_{\text{mri}} = \boldsymbol{I}_{\text{mri}} * L, \ \boldsymbol{H}_{\text{us}} = \boldsymbol{I}_{\text{us}} * L,$$

where L is the 3×3 Laplacian filter. Then, the local average of the absolute value of these high-pass images is used to construct the saliency maps:

$$S_{\text{mri}} = |\boldsymbol{H}_{\text{mri}}| * g_{r_g,\sigma_g},\ S_{\text{us}} = |\boldsymbol{H}_{\text{us}}| * g_{r_g,\sigma_g},$$

with g a Gaussian low-pass filter of size $(2r_g + 1) \times (2r_g + 1)$. Subsequently, the weights are determined by constructing the following indicator:

$$P^k_{\text{mri}} = \begin{cases} 1 \text{ if } S^k_{\text{mri}} = \max\left(S^k_{\text{mri}}, S^k_{\text{us}}\right), \\ 0 \text{ otherwise,} \end{cases}$$

where S^k is the saliency value of the pixel k. The weight maps often exhibit noise that may produce artifacts to the fused image. This issue is usually addressed by incorporating spatial consistency, i.e., adjacent pixels with similar brightness or color should have comparable weights. This can be achieved by using guided filtering. Referring to (4.1) and (4.3), it becomes apparent that when the local variance at a position is very small, which means that the pixel is in a flat region in the guidance image, a_k tends to 0 and the filtering output tends to $\bar{P}_k$, i.e., the average of adjacent input pixels. In contrast, if the local variance at a pixel is significant, suggesting an edge region, a_k becomes far from 0, so only the weights on one side of the edge are averaged. In both scenarios, pixels with similar brightness will have similar weights. Guided filtering is performed on the weight map P_{mri} with the source image I_{mri} serving as the guidance image:

$$W^B_{\text{mri}} = \text{GF}_{r1,\epsilon_1}(P_{\text{mri}}, \boldsymbol{I}_{\text{mri}}),\ W^D_{\text{mri}} = \text{GF}_{r2,\epsilon_2}(P_{\text{mri}}, \boldsymbol{I}_{\text{mri}}), \tag{4.8}$$

where r_1, ϵ_1, r_2, and ϵ_2 are the parameters of the guided filter, and W^B_{mri} and W^D_{mri} are the resulting weight maps of the base and detail images. The final results for the MRI are:

$$\bar{\boldsymbol{B}}_{\text{mri}} = W^B_{\text{mri}} \times \boldsymbol{B}_{\text{mri}},\ \ \bar{\boldsymbol{D}}_{\text{mri}} = W^D_{\text{mri}} \times \boldsymbol{D}_{\text{mri}}. \tag{4.9}$$

4.2.4 Weight construction for the US image

The guided filter has shown good results for images corrupted by additive Gaussian noise. However, in practical applications such as US imaging, where the most prominent noise, i.e., speckle, is non-Gaussian, its performance is limited. In US imaging, speckle is mainly caused by constructive and destructive intereferences between scattered US waves. A reasonable trade-off between accuracy and simplicity is to model this speckle as a multiplicative artifact following a Rayleigh distribution [Aysal 2007b].

Recent works on despeckling extended the linear guided filter to a nonlinear filter [Ni 2016], whose expression can be derived using Bayesian Non-Local Means (NLM). We propose to use the results of [Ni 2016] to define the US weights.
To illustrate the derivation of the final closed-form weights, a simplified notation will be used by considering an observed speckled image y and a noise-free image x.

The final weights will be subsequently applied to both the base and detail layers ($\boldsymbol{B}_{\text{us}}$, $\boldsymbol{D}_{\text{us}}$) of the US image. Using these notations, the nonlinear filter is:

$$\hat{x}_i = \sum_j W_{i,j}(y, x^g)\, y_j, \tag{4.10}$$

where $\hat{x}, y, x^g$ are the despeckled image, the speckled observed image, the guidance image, y_j is the jth pixel of y, and $W_{i,j}(y, x^g)$ is a non-linear weight kernel, where i and j are the pixel positions. We propose to define the guidance image using a simple averaging over every patch in the observed image. In [Ni 2016], the modified Bayesian NLM is expressed as:

$$\hat{x}_i = \frac{\sum_{j\in\omega_i} p(\mathbf{y}_i|\mathbf{y}_j,\, \mathbf{x}_i = \mathbf{x}_j,\, \mathbf{x}^{\mathbf{g}})\, p(\mathbf{y}_j,\, \mathbf{x}_i = \mathbf{x}_j,\, \mathbf{x}^{\mathbf{g}})\, y_j}{\sum_{j\in\omega_i} p(\mathbf{y}_i|\mathbf{y}_j,\, \mathbf{x}_i = \mathbf{x}_j,\, \mathbf{x}^{\mathbf{g}})\, p(\mathbf{y}_j,\, \mathbf{x}_i = \mathbf{x}_j,\, \mathbf{x}^{\mathbf{g}})},$$

where $\mathbf{y}_i$ and $\mathbf{y}_j$ are the vectorized patches of size $M \times M$ centered at pixels i and j, $\mathbf{x}_i$ and $\mathbf{x}_j$ are the noise-free patches corresponding to the speckled patches $\mathbf{y}_i$ and $\mathbf{y}_j$, and $\mathbf{x}^{\mathbf{g}}$ is the corresponding patch in the guidance image. For an image with fully developed and independent speckle, the conditional probability is:

$$p(\mathbf{y}_i|\mathbf{y}_j,\, \mathbf{x}_i = \mathbf{x}_j,\, \mathbf{x}^{\mathbf{g}}) = \prod_{m=1}^{M\times M} p(y_{i,m}|y_{j,m},\, x_{i,m} = x_{j,m},\, x^g_m),$$

where $y_{i,m}, y_{j,m}, x_{i,m}, x_{j,m}$ and x^g_m are associated with the mth pixel in the corresponding patches. Without knowledge of $p(y_{i,m}|y_{j,m})$ and assuming the event $y_{i,m}|y_{j,m},\, x_{i,m} = x_{j,m}$ independent on x^g_m, Bayes rule leads to:

$$\begin{aligned} &p(y_{i,m}|y_{j,m},\, x_{i,m} = x_{j,m},\, x^g_m)\, p(y_{j,m},\, x_{i,m} = x_{j,m},\, x^g_m) \\ &\propto \underbrace{p(y_{i,m}, y_{j,m}|x_{i,m} = x_{j,m})}_{\text{likelihood}}\ \underbrace{p(x_{i,m} = x_{j,m}|x^g_m)}_{\text{prior}}\ p(x^g_m), \end{aligned}$$

where the likelihood is the data fidelity term, and the prior is related to the probability of having $x_{i,m} = x_{j,m}$ given the corresponding value x^g_m in the guidance image. Due to the high dimensionality of image patches, the prior distribution $p(x^g_m)$ is assumed to be uniform for simplicity. For US images corrupted by a multiplicative speckle noise distributed according to a Rayleigh distribution, the following probability density function is obtained:

$$p(y_{i,m}|x_{i,m}) = \frac{y_{i,m}}{x^2_{i,m}\sigma^2_\eta} \exp\left(\frac{y^2_{i,m}}{x^2_{i,m}\sigma^2_\eta}\right) I_{]0,+\infty[}(y_{i,m}), \tag{4.11}$$

where $I_{]0,+\infty[}$ is the indicator function on $]0, +\infty[$ and σ^2_η is the noise variance. The conditional density $p(x_{i,m}|x^g_{i,m})$ is also assumed to be Rayleigh with a parameter $\sigma^2_g(x^g_{i,m})^2$. The likelihood and prior introduced before can then be computed as

follows:

$$p(y_{i,m}, y_{j,m}|x_{i,m} = x_{j,m}) = \int_D p(y_{i,m}|\theta)p(y_{j,m}|\theta)d\theta,$$
$$\propto \frac{y_{i,m}y_{j,m}}{(y_{i,m}^2 + y_{j,m}^2)^{\frac{3}{2}}},$$

$$p(x_{i,m} = x_{j,m}|x_m^g) = \exp\left\{-\mathrm{KL}[p(x_{i,m}|x_{i,m}^g), p(x_{j,m}|x_{j,m}^g]\right\},$$

where D is the domain of pixel values and KL is the symmetric Kullback-Leibler divergence between two Rayleigh distributions. Note that the Kullback-Leibler divergence between two Rayleigh distributions of densities f_1 and f_2 with parameters σ_1^2 and σ_2^2 denoted as $\mathrm{KLD}(f_1, f_2)$ is defined by:

$$\mathrm{KLD}(f_1, f_2) = \int_0^\infty \frac{t}{\sigma_1^2} \exp\left(-\frac{t^2}{2\sigma_1^2}\right) \left[\frac{t^2}{2}\left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}\right) + \log\left(\frac{\sigma_2^2}{\sigma_1^2}\right)\right] dt,$$

which can be decomposed as the sum of two integrals

$$I_1 = \log\left(\frac{\sigma_2^2}{\sigma_1^2}\right) \int_0^\infty \frac{t}{\sigma_1^2} \exp\left(-\frac{t^2}{2\sigma_1^2}\right) dt = \log\left(\frac{\sigma_2^2}{\sigma_1^2}\right),$$

and

$$I_2 = \frac{1}{2}\left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}\right) \int_0^\infty \frac{t^3}{\sigma_1^2} \exp(-\frac{t^2}{2\sigma_1^2})dt = \frac{\sigma_1^2}{\sigma_2^2} - 1.$$

The symmetric Kullback-Leibler divergence between the densities f_1 and f_2 is:

$$\mathrm{KL}(f_1, f_2) = \mathrm{KLD}(f_1, f_2) + \mathrm{KLD}(f_2, f_1) = \frac{(\sigma_1^2 - \sigma_2^2)^2}{\sigma_1^2\sigma_2^2}.$$

Therefore,

$$p(x_{i,m} = x_{j,m}|x_m^g) = \exp\left\{-\mathrm{KL}[p(x_{i,m}|x_{i,m}^g), p(x_{j,m}|x_{j,m}^g]\right\},$$
$$\propto \exp\left\{-\frac{\left[(x_{i,m}^g)^2 - (x_{j,m}^g)^2\right]^2}{(x_{i,m}^g)^2(x_{j,m}^g)^2}\right\},$$

The non-linear weights are finally defined as [Deledalle 2009]:

$$W_{i,j}(y, x^g) = \exp\left\{-\frac{1}{K_1}\sum_m \log\left[\frac{(y_{i,m}^2 + y_{j,m}^2)^{\frac{3}{2}}}{y_{i,m}\, y_{j,m}}\right]\right.$$
$$\left. - \frac{1}{K_2}\sum_m \frac{\left[(x_{i,m}^g)^2 - (x_{j,m}^g)^2\right]^2}{(x_{i,m}^g)^2(x_{j,m}^g)^2}\right\}. \tag{4.12}$$

Using the normalization parameters K_1 and K_2 measuring the relative contributions

of y and x^g, one obtains:

$$\bar{x} = \frac{\sum_{j \in \omega_i} W_{i,j}(y, x^g)\, y}{\sum_{j \in \omega_i} W_{i,j}(y, x^g)}. \tag{4.13}$$

This operation is applied to both base and detail US images.

4.3 Experiments

4.3.1 Synthetic data from real MR acquisition

To demonstrate the efficiency of the proposed MR/US fusion method, this section first considers a set of synthetic images. The simulations presented hereafter have been obtained using a real high-resolution MR image intentionally degraded to simulate an image closely resembling that obtained for endometriosis diagnosis. The 3D high-resolution MR volume corresponds to an actual pelvic MRI capturing the uterus, bladder, and endometriosis lesions. A blurred and noisy 3D MRI is then generated from this high-resolution MR volume. More precisely, the HR volume was blurred using a 2D Gaussian filter of standard deviation $\sigma^2 = 4$, and then was contaminated by an additive white Gaussian noise (SNR = 18.17 dB), yielding the MR image displayed in Fig. 4.1(a). A Rayleigh multiplicative noise was applied to the native MR image, yielding the US image in Fig. 4.1(b) (SNR= 4.8 dB). The size of the average filter was set to 31×31 and the guided filter parameters were fixed to their best values by visual inspection of the fused images, leading to $r_1 = 7$, $\epsilon_1 = 1e^{-4}$, $r_2 = 10$, $\epsilon_2 = 1e^{-6}$. The partial normalization parameters were set to $K_1 = 30$ and $K_2 = 40$ as in [Ni 2016]. This fixed parameter setting was used for both datasets.

(a) (b) (c)

Figure 4.1: (a) MR image (b) US image (c) MARIUS image.

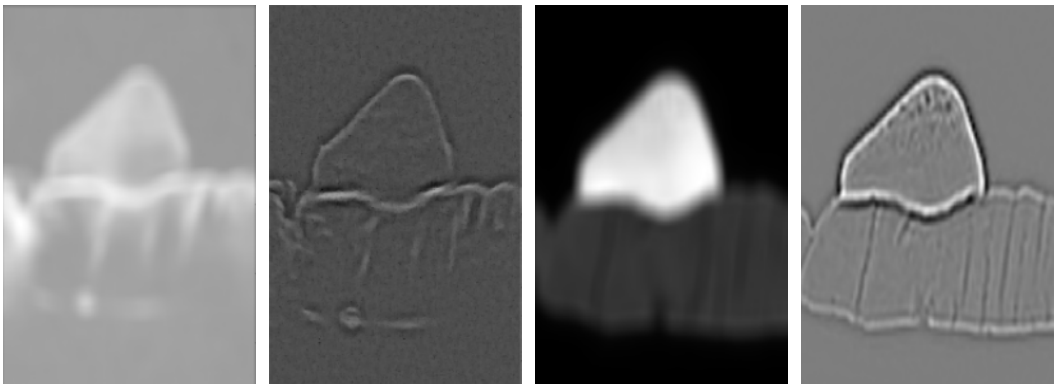
The performance evaluation of the fusion method involved both qualitative analysis through visual inspection of the resultant image and quantitative analysis using the contrast-to-noise ratio (CNR). For two patches extracted from distinct structures,

such as the uterus and bladder in this context, CNR is defined as CNR $= \frac{|\mu_i - \mu_j|}{\sqrt{\sigma_i^2 + \sigma_j^2}}$, where $\mu_i, \mu_j, \sigma_i^2, \sigma_j^2$ are the means and standard deviations of two blocks of pixels. The fused image is displayed in Fig. 4.1(c). This image provides a good compromise between the US and MR data. Specifically, the fused image is less affected by US speckle and MRI blur, provides well-defined contours and good contrast compared to the native MR and US images. The CNR values for the MR and the US images are 35.43 dB and 18.47 dB, whereas the obtained fused image has a CNR equal to 37.54 dB.

4.3.2 Phantom data

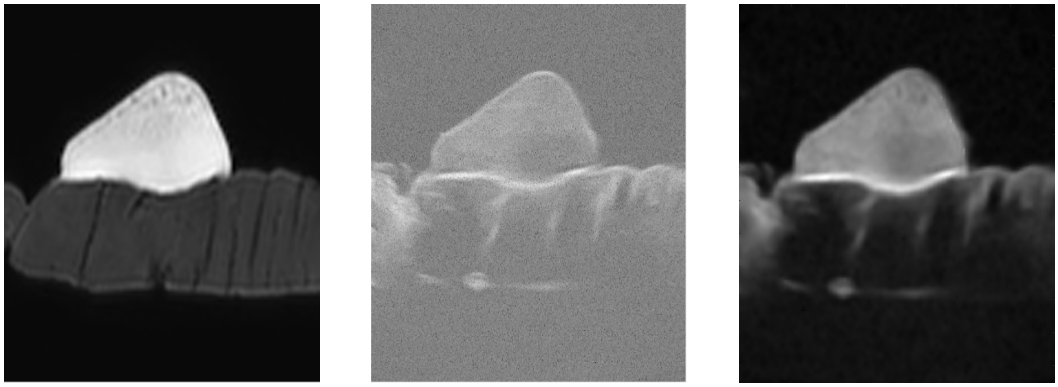
This section evaluates the proposed fusion method on a phantom data. The experimental phantom was designed to replicate the responses of uterine and endometrial tissues to MR and US imaging. It was made of a beef steak on top of which was stuck a polyvinyl alcohol (PVA) phantom, using cyanoacrylate instant glue. On the one hand, the beef meat consists of muscular tissues, exhibiting echogenicity and a response to MR similar to that of uterine tissue. On the other hand, the PVA phantom has roughly the same echogeneicity as the beef meet, but has a different response resulting in high contrast in the MRI. From this viewpoint, its properties are similar to the properties of the endometrium. Finally, the glue between the two structures is visible in US images due to their high resolution but remains absent in MRI due to the more limited resolution, simulating the depth of penetration information, a crucial element for the surgery. Additional details about this experiment can be found in [Vidal 2019].

Fig. 4.2 illustrates the two-scale decomposition of the phantom data as an example, showing both the base and detail layers. The detail layer of the US image emphasizes the contribution to precise contour and edge details, while the base layer of the MRI image highlights the contribution to contrast enhancement.



(a) (b) (c) (d)

Figure 4.2: (a) Base US (b) Detail US (c) Base MRI (d) Detail MRI.



(a) (b) (c)

Figure 4.3: (a) MR image (b) US image (c) MARIUS image.

The fused image of Fig. 4.3 (c) has a spatial resolution similar to the US image, and a contrast equivalent to the MRI. In particular, the different structures of interest are much better highlighted: (i) the glue between the steak and the PVA phantom, simulating the depth of penetration, which is indiscernible in MRI due to resolution limitations, is distinctly visible in both the US and fused images; (ii) the contrast between the steak and the PVA is well defined in the MR and fused images, allowing a clear distinction of the parts of interest.

Conventional fusion methods are then compared to the proposed method. Note that contrary to deep learning (DL) methods, which typically require extensive datasets for training, the proposed approach uses only two images (MRI and US) to perform fusion. Acquiring datasets of registered MR and US images for this specific application is challenging due to privacy concerns and the need for patient consent. Given these constraints, opting for DL methods becomes impractical, reinforcing our decision to benchmark against established non-DL fusion techniques. Quantitivative results are provided in Table 4.1, which compares the CNR values between the PVA phantom and the beef steak for the different images, with fusion models from the literature.

	MRI	US	[El Mansouri 2020]	[Li 2013]	Proposed method
CNR	54.21	18.91	43.17	37.28	47

Table 4.1: CNR (dB) values for the US, MR and fused images.

While the CNR achieved by the fused image may not match that of the MRI, it represents a good compromise between contrast and resolution in the context of endometriosis surgery. The CNR surpasses that of the US, enabling differentiation of various image components. The glue separating the two structures and symbolizing the boundary of the tumor to be incised is more distinctly visible in both the US and fused image compared to the MRI. This improved resolution is visually depicted in Fig. 4.4, illustrating the differences in pixel intensities for the same

vertical straight line (Fig. 4.5) from the US, MR and MARIUS images. For the MRI, the plot shows the change of pixel intensities between the steak and the PVA. For the US, one can see the peak for the glue boundary, that is not highlighted in the MRI due to its low spatial resolution. The fused image shows both the intensity difference and the peak, combining the strengths of both modalities.

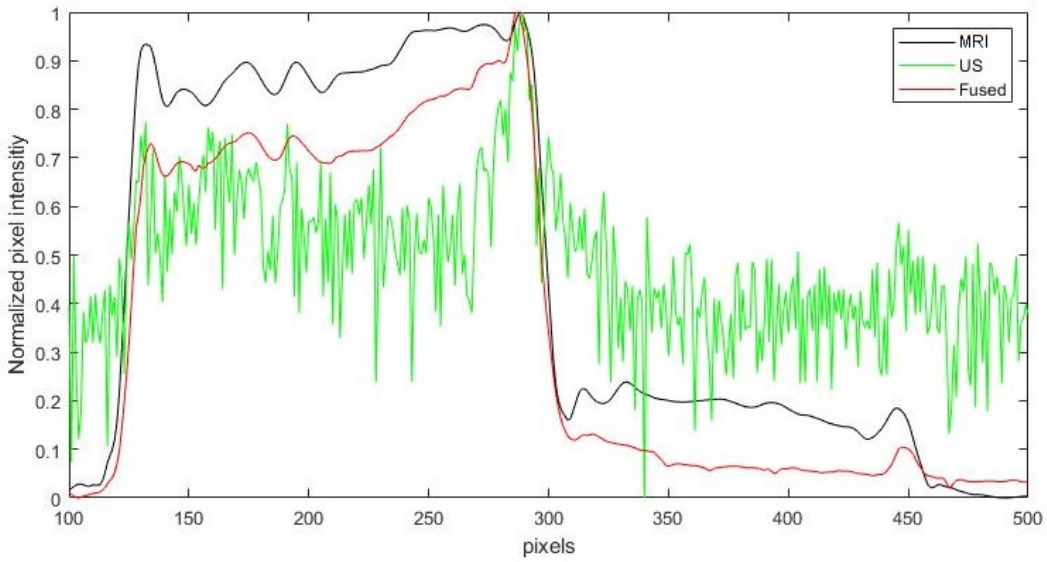


Figure 4.4: Normalized pixel intensities of US (Green), MRI (Black), and fused image (Red) for a vertical straight line.

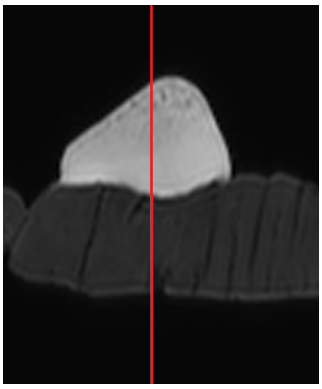


Figure 4.5: The vertical straight line.

This MR/US image fusion based on the guided filtering has demonstrated better performance than existing methods, as evidenced by both quantitative and qualitative results.

Note that the method using kernel transformation, presented in the previous chapter (See section 3.2), achieved a superior CNR on the experimental phantom compared to this model using guided filtering (52.23 dB vs 47 dB). This enhanced

contrast is particularly crucial for applications such as endometriosis imaging, where it helps to distinctly highlight and delineate lesions, improving diagnostic accuracy and facilitating more precise treatment planning. However, it is important to highlight that the kernel transformation-based method requires more computational time to complete the fusion process compared to the guided filtering approach. The extended processing time is a trade-off for the improved contrast quality, which can be a significant consideration in clinical settings where efficiency is also a priority. The guided filtering method, though less complex, still provides satisfactory visual results and is faster because it does not involve calculating coefficients, it may be preferred in situations where processing speed is critical.

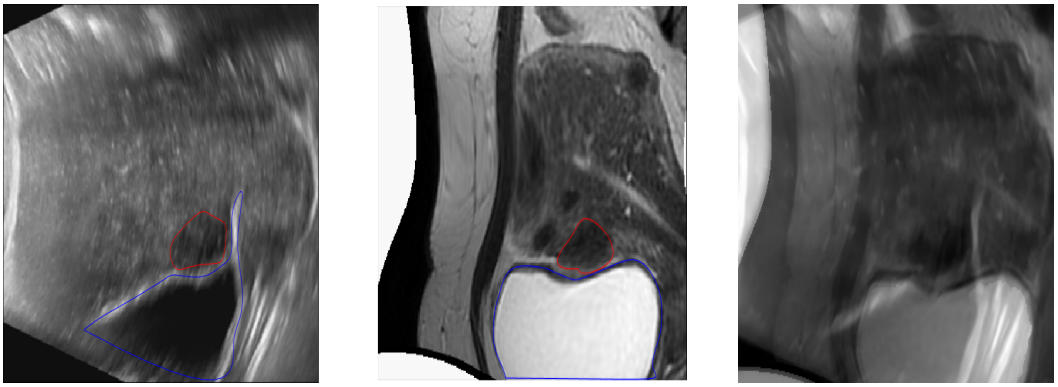
4.3.3 Real data

This section considers experiments using real data to evaluate the guided filtering fusion approach. We use an in vivo MRI volume along with an abdominal US of a patient with endometriosis lesions. US imaging was performed using a Voluson P6 imaging system (GE Healthcare Austria GmbH & Co OG), with a frequency of 7MHz. For the MRI, acquisitions were performed using a 1.5T clinical imaging system (GE Medical Systems Signa). Axial, sagittal, and coronal T2-weighted sequences were obtained with a 4 mm slice thickness and 5 mm spacing between slices, ensuring adequate coverage of the pelvic region. The images were acquired using a 2D acquisition mode with a repetition time (TR) of 3586 ms and an echo time (TE) of 138.0400 ms. The echo train length was set to 34, enhancing the quality of the T2-weighted sequences. The imaging frequency was 63.9053 MHz, targeting the hydrogen nuclei (1H) for optimal signal acquisition.

Registration was performed using the model presented in Section 2.2 with the following parameter setting: maximum translation size $\omega_{trans} = 7$, maximum rotation size $\omega_{rot} = 0.02$, quantization factor $k = 2$, and modification rate $\alpha = 0.8$. The MRI slice corresponding to the US image (Fig. 4.6 (a)) is presented in Fig. 4.6 (b). Fig. 4.6 (c) shows the fusion result. We zoom on the lesion area to better observe the differences between the MRI, US, and fused images. Fig. 4.7 shows the endometriosis lesion in the different images.

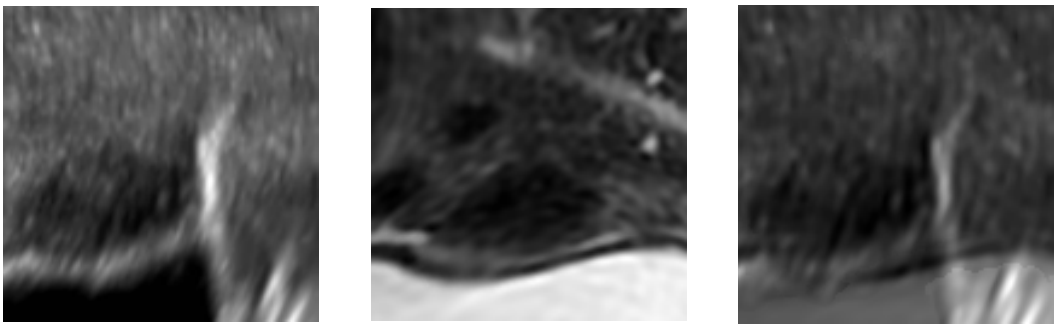
Fig. 4.8 (a) shows pixel intensities of US, MR and the fused image for a vertical straight line (Fig. 4.9). It shows the distinct characteristics of each modality. The MRI exhibits a sudden switch in intensities, highlighting the difference between the bladder and the lesion. The US image, while having a less noticeable difference, shows a peak for the contour line that is indiscernible in the MRI. In the fused image, both the peak and the intensity difference are present, combining the strengths of both modalities.

The proposed fusion method, although gathered information from both input MR and US images, did not ultimately lead to a significant enhancement in image quality. This limitation can be attributed to the nature of the data used in this study. Specifically, the US image available for this research was an abdominal



(a) (b) (c)

Figure 4.6: (a) US image (Endometriosis lesion - Red, Bladder - Blue). (b) MR image. (c) MARIUS image.



(a) (b) (c)

Figure 4.7: (a) Lesion in the US image. (b) Lesion in the MR image. (c) Lesion in the MARIUS image.

ultrasound rather than a transvaginal ultrasound (TVUS). Abdominal ultrasound images typically provide less details compared to TVUS images, which are better suited for pelvic examinations due to their ability to offer higher resolution and more precise anatomical details in the pelvic region. High-frequency ultrasound waves provide higher resolution images but have shorter wavelengths, which limits their depth of penetration. By performing a TVUS, the ultrasound probe is inserted into the vagina, and is much closer to the pelvic organs. This proximity reduces the distance the ultrasound waves need to travel, improving image clarity and detail. For abdominal ultrasound, lower frequencies are used since they can penetrate deeper into the body, allowing for visualization of structures that are further from the surface. However, this comes at the cost of lower resolution compared to TVUS. As a result, the abdominal US image did not contribute as much valuable information as a TVUS image would have.

Given that TVUS images are more effective in visualizing endometrial lesions due to their superior resolution and localized view, the fusion process might have

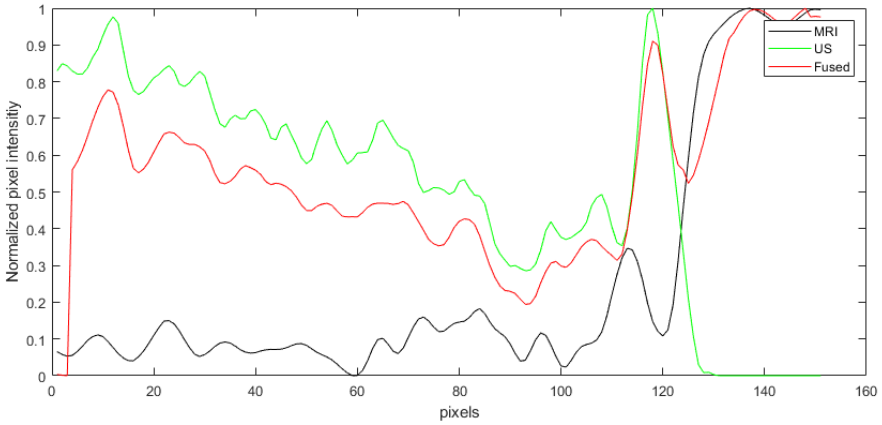


Figure 4.8: Normalized pixel intensities of US (Green), MRI (Black), and fused image (Red) for a vertical straight line.

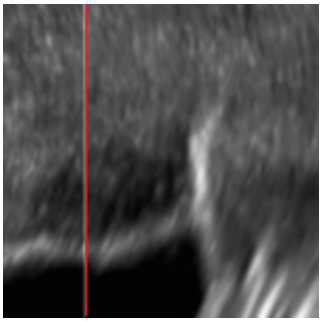


Figure 4.9: The vertical straight line.

yielded better results if TVUS images were used instead. The limitations of the abdominal US image constrained the potential benefit of the fusion method, highlighting the importance of using accurate imaging data to achieve optimal results.

4.4 General framework

In this section, we summarize the general framework explored throughout this PhD thesis. This framework encompasses the comprehensive process of integrating 3D MR and 2D US images to produce the final MARIUS image. The figures below illustrate the complete scheme for both phantom and real data.

The presented figures offer a visual representation of the transition from raw imaging data to the final integrated result. They highlight the potential of combining MR and US imaging modalities to provide more informative and clinically relevant images, ultimately contributing to improved diagnostic accuracy and patient care.

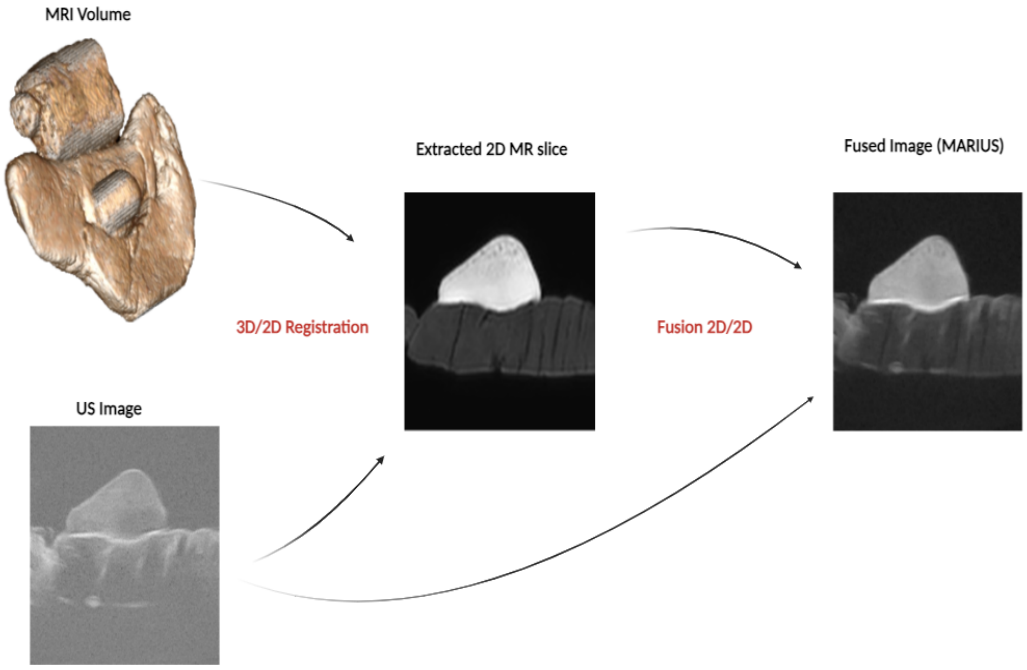


Figure 4.10: Framework for 3D/2D registration and 2D/2D fusion applid to experimental phantom.

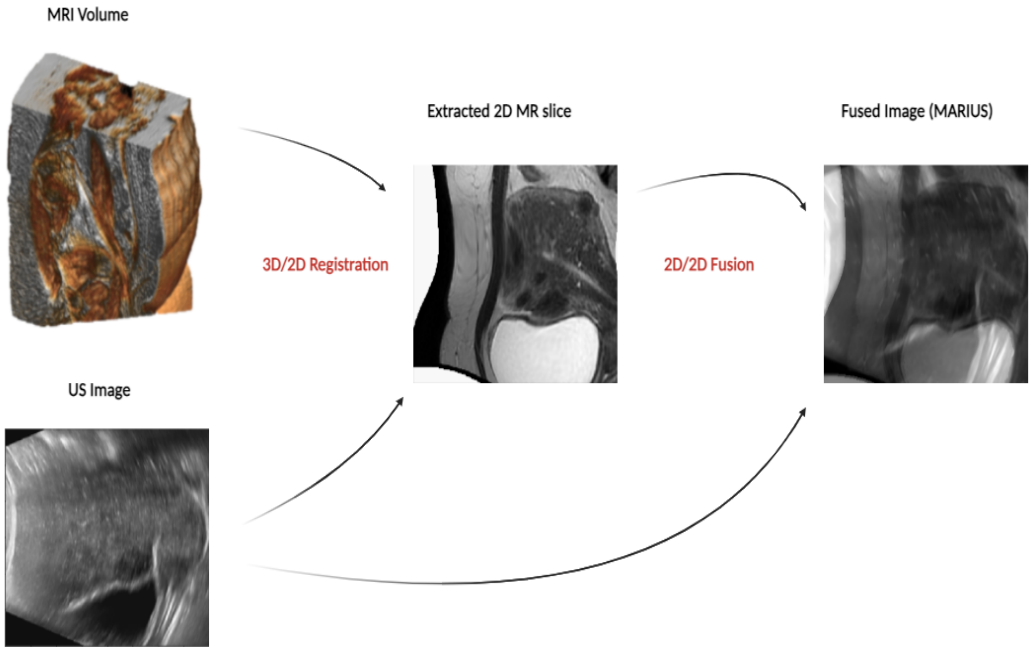


Figure 4.11: Framework for 3D/2D registration and 2D/2D fusion applid to in vivo dataset.

4.5 Conclusion

This chapter introduced a new fusion method for magnetic resonance (MR) and ultrasound (US) images based on guided filtering. The fused image obtained with this method brings together the advantages of both modalities: resolution and contrast. The combination of information arising from both MRI and US into a single image may improve preoperative mapping and surgical plan. An important perspective of this work is to validate the model with other datasets (in vivo data). Future work will be devoted to including the fused image with the video stream collected during laparoscopy, allowing a safer decision-making and therefore a more precise endometriosis surgery.

Chapter 5

Conclusions and Perspectives

Context

Image processing is a powerful tool that has significantly advanced various fields, particularly medical imaging. By enhancing image clarity, contrast, and detail, image processing techniques make it easier for healthcare professionals to interpret and analyze medical images. This, in turn, improves the accuracy of diagnoses, the monitoring of diseases, and the planning and execution of complex medical procedures. Techniques such as image registration and fusion allow for the integration of multiple imaging modalities, providing a more comprehensive view of anatomical structures and potential abnormalities.

Endometriosis is one of the pathologies that requires the use of MR and US modalities in convential clinial practice. This painful condition, which involves the growth of endometrial tissue outside the uterus, can be challenging to diagnose and treat. Precise imaging is crucial for accurately locating and characterizing endometrial lesions. By registring and fusing the images, surgeons can obtain a more detailed and accurate view of the affected areas, allowing for safer decision-making and precise laparoscopies.

The objective of this PhD thesis was to investigate methods for the registration and fusion of magnetic resonance (MR) and ultrasound (US) images, specifically tailored for the pre-operative diagnosis and surgical planning of endometriosis. This task is of sensitive importance as it directly impacts the precision and effectiveness of laparoscopic surgeries. Through this research, we aim to provide robust methodologies that can enhance the visualization of endometrial tissues, leading to more precise and targeted surgical interventions. Experiments were conducted on different datasets, acquired with the collaboration of Dr. Fabien VIDAL (Clinique La Croix du Sud - Ramsay Santé, Toulouse, France). These datasets included synthetic images, an experimental phantom, and a real dataset of a patient with endometriosis. The synthetic images were used as a preliminary step to test our models, allowing us to evaluate their performance under controlled conditions. The experimental phantom, which mimics the response of lesions and tissue to MRI and US imaging, provided a standardized environment to further assess the accuracy and precision of the registration and fusion techniques. Finally, the real dataset of a patient with endometriosis allowed for a practical assessment of the methods in a clinical context.

Conclusions

Chapter 1 presented the clinical context in which this present work takes place. We defined the pathology of endometriosis, detailing its causes, symptoms, and the impact it has on patients. Additionally, we discussed the surgical procedure of laparoscopy, emphasizing its sensitivity and the precision required for successful outcomes. The chapter then introduced the imaging modalities of MRI and US, explaining their advantages and limitations and how these images are formed and used in the diagnosis and treatment of endometriosis.

Chapter 2 gave a general overview of image registration, covering state-of-the-art models for registration and its essential components such as matching criteria, transformation model and optimization algorithms. We then introduced a registration method for 3D MR and 2D US images. The method was able to extract the MR slice best matching the US input image with respect to a chosen similarity measure. The registration took into account the global rigid transformation characterized by rotation and translation parameters, associated with a local deformation based on B-splines. A 2D/2D US/MRI fusion algorithm was also applied to validate the registration results, generating a final image containing the main characteristics of the two images. Results obtained on an experimental phantom, synthetic data generated from an in vivo MRI volume, and a real dataset show the interest of the proposed registration framework.

Chapter 3 presented state-of-the-art models and methods used in image fusion. We then presented a new 2D/2D fusion method for registered MR and US images. The chapter studied the potential interest of modelling the relationship between MR and US images by a non-parametric patch-based transformation defined using the theory of reproducing kernels and evaluate its benefit compared to a global polynomial transformation investigated in the literature. The method is based on an inverse problem, performing a super-resolution of the MR image and a denoising of the US image. A non-linear cost function is used to solve the fusion problem, and its minimization is conducted using PALM algorithm. Qualitative and quantitative results with different datasets allow the performance of the fusion method to be appreciated. However, the proposed appraoch was computationally costly since it may require hundreds of thousands of parameters to be estimated depending on the image and the patch sizes.

In **Chapter 4**, we aimed to explore more efficient alternatives. Motivated by the promising results of guided filtering reported in the literature, we presented a new 2D/2D fusion algorithm for MR and US images. The fused image is a weighted average of base and detail images from the MR and US images. The weights assigned to the MR image allow the contrast of the fused image to be enhanced. For the US image, the key ingredient was to take into account the presence of speckle noise, a common challenge in US imaging. Results on different datasets show very

promising results. The images produced by the proposed fusion method have the potential to enhance visualization and aid decision-making in endometriosis surgery, representing a valuable contribution to the field of medical image fusion.

Future work

Short-term perspectives

- Observation models: For the sake of simplicity and given the complexity and non-linearity of the fusion task, we assumed the presence of independent identically distributed noises in both MRI and US images. A natural progression of this work is to address potential correlations among noise samples, particularly for US images, consider the impact of possible artifacts, and adjust the algorithm for cases of not-fully-developed speckle. Daba et al. developed stochastic models to handle partially-developed speckle noise, where a Poisson point process is modulated by a Gram-Charlier series of Laguerre-weighted exponential functions, resulting in a doubly stochastic filtered Poisson point process [Daba 2008].

- Hyper-parameters: Hyper-parameters play an important role in determining the quality of the obtained fused images (Chapter 3). The choice of these hyper-parameters was done manually, relying on visual inspection of the images and a study of the effect of each parameter on the fused image. We analyzed plots showing the impact of each parameter on appropriate metrics, which guided our selection process. This manual selection process, while effective to some extent, can be subjective and time-consuming. Therefore, it would be interesting to explore methods that can automatically tune these hyper-parameters based on the data. Such automated approaches could enhance the consistency and efficiency of the fusion process, ensuring optimal parameter settings without manual intervention. Future research in this area could significantly improve the robustness and reliability of image fusion techniques [Pereyra 2016].

- Clinical study: The tests conducted in this PhD thesis were first validated using a synthetic dataset with controlled ground truth, then using an experimental phantom designed to mimic the uterus and endometrium responses to MR and US imaging. For the real dataset, we used an in vivo MRI volume and a 2D US abdominal image from a patient with endometriosis. Given the promising results, we are now ready to proceed with a clinical study to further validate and refine the methods developed in this research. This study should include transvaginal US images that are actually used in practice and are better suited for the application. Additionally, exploring the application of the proposed methods to other in vivo datasets for the same purpose would be an interesting direction for future research.

Long-term perspectives

- Augmented reality-assisted surgery: The 3D/2D registration and 2D/2D fusion of MR and US images marks the initial phase of an extensive study aimed at developing a virtual navigator based on augmented reality for endometriosis surgery [Collins 2021] [Shahzad 2023]. Augmented reality is an evolving technology that overlays computer-generated information onto the real-world environment. It transforms the user's field of view into a dynamic display, where real-world objects complement virtual data [Dennler 2021]. For instance, Pratt et. al. [Pratt 2018] used Microsoft HoloLens™ stereo head-mounted display, an augmented reality technology for reconstructive surgery. They used pre-operative computed tomography angiography imaging to help the surgeon see through the patient's skin. Fig. 5.1 shows the patient's body as viewed by the surgeon using HoloLens, where information from pre-operative imaging is imposed.

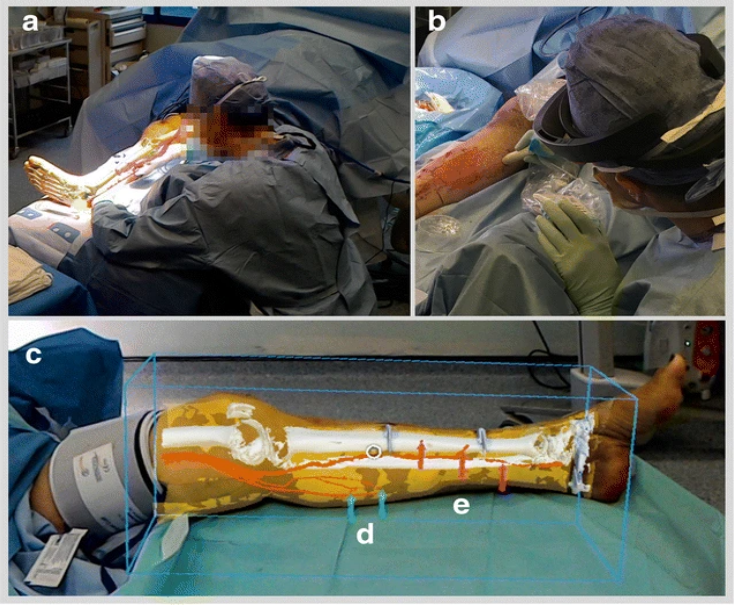


Figure 5.1: Augmented reality overlay of models as viewed by the surgeon from remote HoloLens for reconstructive surgery [Pratt 2018].

For endometriosis surgery, the goal of this future study is to combine this newly fused MARIUS image, which combines the advantages of both MR and US images, into the video stream collected using a small camera during laparoscopy. This integration should provide real-time, detailed anatomical information overlaid directly onto the surgeon's view, improving spatial awareness and accuracy during procedures. Surgeons can benefit from enhanced visualization of the critical structures, precise guidance for instrument placement, and improved decision-making [Qian 2019] [Birlo 2022]. Ultimately, this technology has the potential to make laparoscopies safer, more effective,

reducing surgical complications and improving the lives of many women.

- Other medical applications: Many other medical applications use MR and US images and could benefit from the proposed registration and fusion framework. Registration and fusion of mammograms with breast MRI or ultrasound can provide a more comprehensive view of the tumor's extent and its relationship with surrounding tissues [Bessa 2020] [Sivaramakrishna 2005]. For liver tumors, fusion of preoperative imaging (such as MRI or CT) with real-time ultrasound or intraoperative imaging allows for accurate targeting of the tumor. This ensures that the ablation is applied precisely to the tumor while minimizing damage to surrounding healthy tissue. After the ablation [Mauri 2014], fused images can be used to assess the treatment's success by comparing pre- and post-procedure images. This helps in determining if the tumor has been adequately treated or if further intervention is required [Spinczyk 2020]. Other applications include Prostate biopsy [Gao 2013] Cardiology and [Tavard 2014].

Appendix A

Update of x in the PALM algorithm of Section 3.2.4.1

A.1 Lipshitz constant

The Lipshitz constant $L_{\boldsymbol{x}}(\boldsymbol{v})$ of $\boldsymbol{x} \mapsto \nabla_{\boldsymbol{x}} H(\boldsymbol{x}, \boldsymbol{v})$ is complicated to evaluate and not straighforward. In [El Mansouri 2020], PALM with a backtracking stepsize rule was used:

Note that one of the pillars of PALM's convergence proof is the following lemma for smooth functions.

Lemma: Let $(h : \mathbb{R}^n \to \mathbb{R})$ be a continuously differentiable function with Lipschitz continuous gradient and Lipschitz constant L_h. Then for any $L \geq L_h$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$,

$$h(\boldsymbol{x}) \leq h(\boldsymbol{y}) + (\boldsymbol{x} - \boldsymbol{y})^H \nabla_x h(\boldsymbol{x}) + \frac{L}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2.$$

Using simple algebra, one can show that the function ψ defined in 3.16 satisfies the following relation:

$$\psi(\boldsymbol{x}, \boldsymbol{v}) \leq Q_L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{v}), \tag{A.1}$$

where

$$Q_L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{v}) = l(\boldsymbol{x}) + g(\boldsymbol{v}) + H(\boldsymbol{y}, \boldsymbol{v}) + (\boldsymbol{x} - \boldsymbol{y})^H \nabla_{\boldsymbol{x}} H(\boldsymbol{y}, \boldsymbol{v}) + \frac{L}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2.$$

In order to ensure the convergence of the PALM algorithm, the backtracking rule consists of verifying that the inequality A.1 is satisfied at every step. To estimate the Lipschitz constant, at each iteration k, we search for the smallest nonnegative integer i_k such that $L_{k+1} = \lambda^{i_k} L_k$ verifies

$$\psi(p_{\bar{L}}(\boldsymbol{x}^k), \boldsymbol{v}^k) \leq Q_{\bar{L}}(p_{\bar{L}}(\boldsymbol{x}^k), \boldsymbol{x}^k, \boldsymbol{v}^k),$$

with

$$p_{\bar{L}}(\boldsymbol{x}^k) = \text{prox}^l_{L_{k+1}}\left(\boldsymbol{x}^k - \frac{1}{L_{k+1}} \nabla_{\boldsymbol{x}} H(\boldsymbol{x}^k, \boldsymbol{v}^k)\right).$$

Remark: Since inequality A.1 is satisfied for $L \geq L_{\boldsymbol{x}}(\boldsymbol{v})$, for all $\boldsymbol{v}$, where

$L_{\boldsymbol{x}}(\boldsymbol{v})$ is the Lipschitz constant of $x \mapsto \nabla_{\boldsymbol{x}} H(\boldsymbol{x}, \boldsymbol{v})$, the following inequalities can be obtained:

$$L_x(\boldsymbol{v}^k) \leq L_{k+1} \leq \lambda L_H(\boldsymbol{v}^k),$$

where $L_H(\boldsymbol{v}^k)$ denotes the Lipschitz constant of H with respect to v at $\boldsymbol{v}^k$. Note that the inequality $L_x(v_k) \leq L_{k+1}$ is sufficient to ensure the convergence of PALM. However, the second inequality $L_{k+1} \leq \lambda L_H(\boldsymbol{v}^k)$ allows the convergence rate to be controlled by an appropriate choice of λ.

Finally, once the smallest nonnegative integers i_k such that

$$\psi(p_{\bar{L}}(\boldsymbol{x}^k), \boldsymbol{v}^k) \leq Q_{\bar{L}}(p_{\bar{L}}(\boldsymbol{x}^k), \boldsymbol{x}^k, \boldsymbol{v}^k),$$

with $\bar{L} = i_k L_k$ is found, we set L_{k+1} to $L_{k+1} = i_k L_k$ and $\boldsymbol{x}$ is updated.

A.2 Update of $\boldsymbol{x}$

The update of $\boldsymbol{x}$ is done using Eq. 3.20. Since it requires the inversion of a high-dimensional matrix, we adopt the solution used in [El Mansouri 2020]. The blurring matrix $\boldsymbol{C}$ is block circulant and can be diagonalized in the Fourier domain:

$$\boldsymbol{C} = \boldsymbol{F}^H \boldsymbol{\Lambda} \boldsymbol{F},$$

where $\boldsymbol{F}$ and $\boldsymbol{F}^H$ are the 2D Fourier and inverse Fourier operators, $\boldsymbol{\Lambda} = \text{diag}(\boldsymbol{F}\boldsymbol{h})$, with $\boldsymbol{h}$ the first column of $\boldsymbol{C}$. The equation (3.20) can thus be rewritten as follows:

$$\boldsymbol{x}^{k+1} = \frac{1}{\tau_1 L_{k+1}} \boldsymbol{F}^H \boldsymbol{\psi} \boldsymbol{F} \boldsymbol{r} - \frac{1}{\tau_1 L_{k+1}} \boldsymbol{F}^H \boldsymbol{\psi} \underline{\boldsymbol{\Lambda}}^H (2\tau_1 s \boldsymbol{I}_N + \underline{\boldsymbol{\Lambda}} \boldsymbol{\psi} \boldsymbol{\Lambda}^H)^{-1} \underline{\boldsymbol{\Lambda}} \boldsymbol{\psi} \boldsymbol{F} \boldsymbol{r},$$

with

$$\boldsymbol{r} = \boldsymbol{C}^H \boldsymbol{S}^H \boldsymbol{y}_{\text{mri}} + L_{k+1}(\boldsymbol{x}^k - \frac{1}{L_{k+1}} \nabla_{\boldsymbol{x}} H(\boldsymbol{x}^k, \boldsymbol{v}^k))$$

and

$$\underline{\boldsymbol{\Lambda}} = \text{diag}(\boldsymbol{\Lambda_1}, ..., \boldsymbol{\Lambda_d})$$

is a block diagonal matrix and its diagonal elements are matrices $\boldsymbol{\Lambda_i}$, and where

$$\boldsymbol{\psi} = \boldsymbol{F} \left(\tau_1 \left(\boldsymbol{D}_h^H \boldsymbol{D}_h + \boldsymbol{D}_v^H \boldsymbol{D}_v \right) + \frac{L_{k+1}}{2} I_N \right)^{-1} \boldsymbol{F}^H$$

Bibliography

[Al-Majeed 2020] Al-Majeed, S. and Al-Tamimi, M. *Survey Based Study: Classification of Patients with Alzheimer's Disease, volume = 61, journal = Iraqi Journal of Science, doi = 10.24996/ijs.2020.61.11.31.* pages 3104–3126, 12 2020. (Cited in pages ix and 16.)

[Amberg 2007] Amberg, B., Romdhani, S. and Vetter, T. *Optimal Step Nonrigid ICP Algorithms for Surface Registration.* In 2007 IEEE Conference on Computer Vision and Pattern Recognition, pages 1–8, 2007. (Cited in page 31.)

[Amin-Naji 2018] Amin-Naji, M. and Aghagolzadeh, A. *Multi-focus image fusion in DCT domain using variance and energy of Laplacian and correlation coefficient for visual sensor networks.* Journal of AI and Data Mining, vol. 6, no. 2, pages 233–250, 2018. (Cited in page 60.)

[Anila Satheesh 2021] Anila Satheesh, B. and Thittai, A. K. *Ultrasound-to-Ultrasound deformable image registration for pseudo-CT simulation for Brachytherapy treatment planning: A preliminary study on tissue-mimicking phantom.* In TENCON 2021 - 2021 IEEE Region 10 Conference (TENCON), pages 901–904, 2021. (Cited in page 26.)

[Atehortúa 2019] Atehortúa, A., Garreau, M., Simon, A., Donal, E., Lederlin, M. and Romero, E. *Fusion of 3D real-time echocardiography and cine MRI using a saliency analysis.* International Journal of Computer Assisted Radiology and Surgery, vol. 15, 11 2019. [Online]. Available: http://dx.doi.org/10.1007/s11548-019-02087-w. (Cited in pages 26 and 36.)

[Aysal 2007a] Aysal, T. and Barner, K. E. *Rayleigh-maximum-likelihood filtering for speckle reduction of ultrasound images.* IEEE Trans. Med. Imag., vol. 26, no. 5, pages 712–727, 2007. (Cited in pages 60 and 69.)

[Aysal 2007b] Aysal, T. C. and Barner, K. E. *Rayleigh-Maximum-Likelihood Filtering for Speckle Reduction of Ultrasound Images.* IEEE Transactions on Medical Imaging, vol. 26, no. 5, pages 712–727, 2007. [Online]. Available: http://dx.doi.org/10.1109/TMI.2007.895484. (Cited in page 86.)

[Barbieri 1982] Barbieri, R. L., Evans, S. and Kistner, R. W. *Danazol in the treatment of endometriosis: analysis of 100 cases with a 4-year follow-up.* Fertility and Sterility, vol. 37, no. 6, pages 737–746, 1982. (Cited in page 5.)

[Bennioui 2023] Bennioui, Y. E., Vidal, F., Basarab, A. and Tourneret, J.-Y. *Joint Registration and Fusion of 3D Magnetic Resonance and 2D Ultrasound Images for Endometriosis Surgery.* In 2023 31st European Signal Processing Conference (EUSIPCO), pages 1130–1134, 2023. (Cited in pages 21 and 84.)

[Bentoutou 2005] Bentoutou, Y., Taleb, N., Kpalma, K. and Ronsin, J. *An automatic image registration for applications in remote sensing.* IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 9, pages 2127–2137, 2005. [Online]. Available: http://dx.doi.org/10.1109/TGRS.2005.853187. (Cited in page 26.)

[Besl 1992] Besl, P. and McKay, N. D. *A method for registration of 3-D shapes.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14, no. 2, pages 239–256, 1992. [Online]. Available: http://dx.doi.org/10.1109/34.121791. (Cited in page 29.)

[Bessa 2020] Bessa, S., Gouveia, P. F., Carvalho, P. H., Rodrigues, C., Silva, N. L., Cardoso, F., Cardoso, J. S., Oliveira, H. P. and Cardoso, M. J. *3D digital breast cancer models with multimodal fusion algorithms.* The Breast, vol. 49, pages 281–290, 2020. (Cited in page 104.)

[Bhavana 2015] Bhavana, V. and H.K., K. *Multi-Modality Medical Image Fusion using Discrete Wavelet Transform.* Procedia Computer Science, vol. 70, pages 625–631, 2015. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1877050915032214`. Proceedings of the 4th International Conference on Eco-friendly Computing and Communication Systems. (Cited in pages 60 and 67.)

[Birlo 2022] Birlo, M., Edwards, P. E., Clarkson, M. and Stoyanov, D. *Utility of optical see-through head mounted displays in augmented reality-assisted surgery: A systematic review.* Medical Image Analysis, vol. 77, page 102361, 2022. (Cited in page 103.)

[Blasch 2008] Blasch, E., Li, X., Chen, G. and Li, W. *Image quality assessment for performance evaluation of image fusion.* In 2008 11th International Conference on Information Fusion, pages 1–6, 2008. (Cited in page 67.)

[Bolte 2014] Bolte, J., Shoham, S. and Teboulle, M. *Proximal alternating linearized minimization or nonconvex and nonsmooth problems.* Mathematical Programming, vol. 146, no. 1-2, pages 459–494, 2014. (Cited in pages 71 and 72.)

[Bouboulis 2010] Bouboulis, P., Slavakis, K. and Theodoridis, S. *Adaptative Kernel-based Image Denoising employing Semi-Parametric Regularization.* IEEE Trans. Image Process., vol. 19, no. 1, pages 1465–79, 2010. (Cited in page 71.)

[Brown 2017] Brown, J., Crawford, T. J., Allen, C., Hopewell, S. and Prentice, A. *Nonsteroidal anti-inflammatory drugs for pain in women with endometriosis.* Cochrane Database of Systematic Reviews, no. 1, 2017. (Cited in page 5.)

[Bulletti 2010] Bulletti, C., Coccia, M. E., Battistoni, S. and Borini, A. *Endometriosis and infertility.* Journal of assisted reproduction and genetics, vol. 27, pages 441–447, 2010. (Cited in page 2.)

[Byun 2015] Byun, Y., Han, Y. and Chae, T. *Image fusion-based change detection for flood extent extraction using bi-temporal very high-resolution satellite images.* Remote Sensing, vol. 7, no. 8, pages 10347–10363, 2015. (Cited in page 60.)

[Calhoun 2009] Calhoun, V. D. and Adali, T. *Feature-Based Fusion of Medical Imaging Data.* IEEE Transactions on Information Technology in Biomedicine, vol. 13, no. 5, pages 711–720, 2009. [Online]. Available: http://dx.doi.org/10.1109/TITB.2008.923773. (Cited in page 61.)

[Chan 2011] Chan, V. and Perlas, A. *Basics of ultrasound imaging.* Atlas of ultrasound-guided procedures in interventional pain management, pages 13–19, 2011. (Cited in page 8.)

[Chui 2003] Chui, H. and Rangarajan, A. *A new point matching algorithm for non-rigid registration.* Computer Vision and Image Understanding, vol. 89, no. 2, pages 114–141, 2003. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1077314203000092`. Nonrigid Image Registration. (Cited in page 32.)

[Collins 2021] Collins, T., Pizarro, D., Gasparini, S., Bourdel, N., Chauvet, P., Canis, M., Calvet, L. and Bartoli, A. *Augmented Reality Guided Laparoscopic Surgery of the Uterus.* IEEE Transactions on Medical Imaging, vol. 40, no. 1, pages 371–380, 2021. [Online]. Available: http://dx.doi.org/10.1109/TMI.2020.3027442. (Cited in page 103.)

[Craig 2020] Craig, E. V., Shannon, L. M. and Andreotti, R. F. *The complementary role of ultrasound and magnetic resonance imaging in the evaluation of endometriosis: a review.* Ultrasound Quarterly, vol. 36, no. 2, pages 123–132, 2020. (Cited in page 6.)

[Daba 2008] Daba, J. and Jreije, P. *Advanced Stochastic Models for Partially Developed Speckle.* vol. 2, pages 1001–1005, 02 2008. (Cited in page 102.)

[Daniilidis 2022] Daniilidis, A., Grigoriadis, G., Dalakoura, D., D'Alterio, M. N., Angioni, S. and Roman, H. *Transvaginal Ultrasound in the Diagnosis and Assessment of Endometriosis—An Overview: How, Why, and When.* Diagnostics, vol. 12, no. 12, page 2912, 2022. (Cited in pages ix and 17.)

[Darrow 1993] Darrow, S. L., Vena, J. E., Batt, R. E., Zielezny, M. A., Michalek, A. M. and Selman, S. *Menstrual cycle characteristics and the risk of endometriosis.* Epidemiology, vol. 4, no. 2, pages 135–142, 1993. (Cited in page 4.)

[De Craene 2004] De Craene, M., du Bois d'Aische, A., Macq, B., Kipfmueller, F., Weisenfeld, N., Haker, S. and Warfield, S. K. *Multimodal nonrigid registration using a stochastic gradient approximation.* In 2004 2nd IEEE Interna-

tional Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821), pages 1459–1462. IEEE, 2004. (Cited in page 36.)

[Deledalle 2009] Deledalle, C.-A., Denis, L. and Tupin, F. *Iterative Weighted Maximum Likelihood Denoising With Probabilistic Patch-Based Weights.* IEEE Transactions on Image Processing, vol. 18, no. 12, pages 2661–2672, 2009. [Online]. Available: http://dx.doi.org/10.1109/TIP.2009.2029593. (Cited in page 88.)

[Deng 2016] Deng, L.-J., Guo, W. and Huang, T.-Z. *Single-Image Super-Resolution via an Iterative Reproducing Kernel Hilbert Space Method.* IEEE Trans. Circuits Syst. Video Technol., vol. 26, no. 11, pages 2001–2014, 2016. (Cited in page 71.)

[Dennler 2021] Dennler, C., Bauer, D., Scheibler, A.-G., Spirig, J., Götschi, T., Fürnstahl, P. and Farshad, M. *Augmented reality in the operating room: a clinical feasibility study.* BMC Musculoskeletal Disorders, vol. 22, 05 2021. [Online]. Available: http://dx.doi.org/10.1186/s12891-021-04339-w. (Cited in page 103.)

[Denny 2007] Denny, E. and Mann, C. H. *Endometriosis-associated dyspareunia: the impact on women's lives.* BMJ Sexual & Reproductive Health, vol. 33, no. 3, pages 189–193, 2007. (Cited in page 2.)

[Dhannawat 2013] Dhannawat, R. *"Kekre's Hybrid Wavelet Transform technique with DCT, Walsh, Hartely and Kekre's transform for Image Fusion" in International Journal of Computer Engineering and Technology (IJCET), Vol. 4, No. 1, Feb 2013, page(s): 195- 202.* International Journal of Computer Engineering and Technology (IJCET), vol. 4, pages 195–202, 02 2013. (Cited in page 64.)

[Dréo 2006] Dréo, J., Nunes, J.-C. and Siarry, P. *Robust rigid registration of retinal angiograms through optimization.* Computerized Medical Imaging and Graphics, vol. 30, no. 8, pages 453–463, 2006. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0895611106000760`. (Cited in page 35.)

[Duffy 2014] Duffy, J. M., Arambage, K., Correa, F. J., Olive, D., Farquhar, C., Garry, R., Barlow, D. H. and Jacobson, T. Z. *Laparoscopic surgery for endometriosis.* Cochrane Database of Systematic Reviews, no. 4, 2014. (Cited in page 5.)

[El Bennioui 2023a] El Bennioui, Y., Bruguier, A., Vidal, F., Basarabr, A. and Tourneret, J.-Y. *Fusion of Ultrasound and Magnetic Resonance Images for Endometriosis Diagnosis: A Non-Parametric Approach.* In 2023 IEEE 9th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), pages 26–30. IEEE, 2023. (Cited in page 21.)

[El Bennioui 2023b] El Bennioui, Y., Mansouri, O., Vidal, F., Basarab, A. and Tourneret, J.-Y. *Recalage et fusion d'images échographiques et par résonance magnétique au service de la chirurgie de l'endométriose.* 08 2023. (Cited in page 21.)

[El Bennioui 2024] El Bennioui, Y., Halimi, A., Basarab, A. and Tourneret, J.-Y. *Fusion of Magnetic Resonance and Ultrasound Images Using Guided Filtering: Application to Endometriosis Surgery.* In 32nd European conference on signal processing 2024, 2024. (Cited in page 22.)

[El Mansouri 2020] El Mansouri, O., Vidal, F., Basarab, A., Payoux, P., Kouamé, D. and Tourneret, J.-Y. *Fusion of Magnetic Resonance and Ultrasound Images for Endometriosis Detection.* IEEE Transactions on Image Processing, vol. 29, pages 5324–5335, 2020. [Online]. Available: http://dx.doi.org/10.1109/TIP.2020.2975977. (Cited in pages x, 19, 21, 44, 48, 52, 59, 60, 69, 71, 74, 75, 77, 78, 82, 91, 105, and 106.)

[Elen 2010] Elen, A., Hermans, J., Ganame, J., Loeckx, D., Bogaert, J., Maes, F. and Suetens, P. *Automatic 3-D Breath-Hold Related Motion Correction of Dynamic Multislice MRI.* IEEE Transactions on Medical Imaging, vol. 29, no. 3, pages 868–878, 2010. [Online]. Available: http://dx.doi.org/10.1109/TMI.2009.2039145. (Cited in page 34.)

[Endometriosis-UK] Endometriosis-UK. *What is endometriosis?* `https://www.endometriosis-uk.org/what-endometriosis`. (Cited in pages ix and 2.)

[Eskenazi 1997] Eskenazi, B. and Warner, M. L. *Epidemiology of endometriosis.* Obstetrics and gynecology clinics of North America, vol. 24, no. 2, pages 235–258, 1997. (Cited in page 1.)

[Fallas 1956] Fallas, R. E. *Endometriosis demonstration for the Sampson theory by a human anomaly.* American Journal of Obstetrics and Gynecology, vol. 72, no. 3, pages 557–561, 1956. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/0002937856903805`. (Cited in page 4.)

[Farnia 2012] Farnia, P., Ahmadian, A., Sedighpoor, M., Khoshnevisan, A. and Siyah Mansoory, M. *On the performance of improved ICP algorithms for registration of intra-ultrasound with pre-MR images; a phantom study.* In 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pages 4390–4393, 2012. (Cited in page 29.)

[Fazleabas 2002] Fazleabas, A. T., Brudney, A., Gurates, B., Chai, D. and Bulun, S. *A modified baboon model for endometriosis.* Annals of the New York Academy of Sciences, vol. 955, no. 1, pages 308–317, 2002. (Cited in pages 1 and 3.)

[Fei 2003] Fei, B., Frinkley, K. and Wilson, D. *Registration algorithms for interventional MRI-guided treatment of the prostate.* Proc SPIE, vol. 5029, pages

192–201, 05 2003. [Online]. Available: http://dx.doi.org/10.1117/12.479696. (Cited in page 35.)

[Fei 2004] Fei, B., Lee, Z., Boll, D., Duerk, J., Sodee, D., Lewin, J. and Wilson, D. *Registration and Fusion of SPECT, High-Resolution MRI, and Interventional MRI for Thermal Ablation of Prostate Cancer.* Nuclear Science, IEEE Transactions on, vol. 51, pages 177 – 183, 03 2004. [Online]. Available: http://dx.doi.org/10.1109/TNS.2003.823027. (Cited in page 37.)

[Ferrante 2013] Ferrante, E. and Paragios, N. *Non-rigid 2D-3D Medical Image Registration Using Markov Random Fields.* In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013, pages 163–170, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg. (Cited in page 28.)

[Ferrante 2015] Ferrante, E., Fecamp, V. and Paragios, N. *Implicit planar and in-plane deformable mapping in medical images through high order graphs.* In 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), pages 721–724, 2015. (Cited in page 33.)

[Ferrante 2017] Ferrante, E. and Paragios, N. *Slice-to-volume medical image registration: A survey.* Medical Image Analysis, vol. 39, pages 101–123, 2017. (Cited in page 37.)

[Fertility 2024] Fertility, S. *Treatment of endometriosis*, 2024. `https://www.safertility.co.uk/Treatment_of_endometriosis`. (Cited in pages ix and 3.)

[Fogtmann 2014] Fogtmann, M., Seshamani, S., Kroenke, C., Cheng, X., Chapman, T., Wilm, J., Rousseau, F. and Studholme, C. *A Unified Approach to Diffusion Direction Sensitive Slice Registration and 3-D DTI Reconstruction From Moving Fetal Brain Anatomy.* IEEE Trans. Med. Imag., vol. 33, no. 2, 2014. [Online]. Available: http://dx.doi.org/10.1109/TMI.2013.2284014. (Cited in page 28.)

[Fordham 2021] Fordham, A.-J., Hacherl, C.-C., Patel, N., Jones, K., Myers, B., Abraham, M. and Gendreau, J. *Differentiating Glioblastomas from Solitary Brain Metastases: An Update on the Current Literature of Advanced Imaging Modalities.* Cancers, vol. 13, page 2960, 06 2021. [Online]. Available: http://dx.doi.org/10.3390/cancers13122960. (Cited in pages ix and 14.)

[Fuerst 2014] Fuerst, B., Wein, W., Müller, M. and Navab, N. *Automatic ultrasound–MRI registration for neurosurgery using the 2D and 3D LC2 Metric.* Medical Image Analysis, vol. 18, no. 8, pages 1312–1319, 2014. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1361841514000620`. Special Issue on the 2013 Conference on Medical Image Computing and Computer Assisted Intervention. (Cited in page 33.)

[Gao 2013] Gao, J., Zhou, F. and Liu, B. *Elastic image registration of MR/TRUS for targeted prostate biopsy.* In 2013 IEEE International Conference on

Medical Imaging Physics and Engineering, pages 245–249, 2013. (Cited in page 104.)

[Garg 2005] Garg, S., Kiran, K., Mohan, R. and Tiwary, U. *Multilevel Medical Image Fusion using Segmented Image by Level Set Evolution with Region Competition.* In 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, pages 7680–7683, 2005. (Cited in page 64.)

[Ge 2008] Ge, W., Yuan, G., Li, C., Wu, Y., Zhang, Y. and Xu, X. *CT image fusion in the evaluation of radiation treatment planning for non-small cell lung cancer.* The Chinese-German Journal of Clinical Oncology, vol. 7, pages 315–318, 2008. (Cited in page 60.)

[Geng 2019] Geng, P., Wang, Z. and Sun, X. *Image fusion adopting guided filter and multi-resolution singular value decomposition.* Journal of Information Hiding and Multimedia Signal Processing, vol. 10, pages 19–29, 01 2019. (Cited in page 66.)

[Genovese 2016] Genovese, M. *Ultrasound transducers.* Journal of Diagnostic Medical Sonography, vol. 32, no. 1, pages 48–53, 2016. (Cited in page 7.)

[Ghai 2020] Ghai, V., Jan, H., Shakir, F., Haines, P. and Kent, A. *Diagnostic delay for superficial and deep endometriosis in the United Kingdom.* Journal of Obstetrics and Gynaecology, vol. 40, no. 1, pages 83–89, 2020. (Cited in page 3.)

[Gholipour 2010] Gholipour, A., Estroff, J. A. and Warfield, S. K. *Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain MRI.* IEEE Trans. Med. Imag., vol. 29, no. 10, pages 1739–1758, 2010. (Cited in pages 59 and 69.)

[Giesel 2009] Giesel, F., Mehndiratta, A., Locklin, J., McAuliffe, M., White, S., Choyke, P., Knopp, M., Wood, B., Haberkorn, U. and von Tengg-Kobligk, H. *Image fusion using CT, MRI and PET for treatment planning, navigation and follow up in percutaneous RFA.* Experimental oncology, vol. 31, no. 2, page 106, 2009. (Cited in page 60.)

[Gonzalez 2008] Gonzalez, R. C. and Woods, R. E. Digital image processing. Prentice Hall, Upper Saddle River, N.J., 2008. (Cited in page 85.)

[Gore 2003] Gore, J. C.*et al. Principles and practice of functional MRI of the human brain.* The Journal of clinical investigation, vol. 112, no. 1, pages 4–9, 2003. (Cited in page 16.)

[Goreczny 2017] Goreczny, S., Moszura, T., Dryzek, P., Lukaszewski, M., Krawczuk, A., Moll, J. and Morgan, G. *Three-dimensional image fusion guidance of percutaneous pulmonary valve implantation to reduce radiation exposure and contrast dose: a comparison with traditional two-dimensional*

and three-dimensional rotational angiographic guidance. Netherlands Heart Journal, vol. 25, pages 91–99, 2017. (Cited in page 60.)

[Greenspan 2002] Greenspan, H., Oz, G., Kiryati, N. and Peled, S. *MRI inter-slice reconstruction using super-resolution.* Magnetic resonance imaging, vol. 20, no. 5, pages 437–446, 2002. (Cited in pages 59 and 69.)

[Guerriero 2016] Guerriero, S., Ajossa, S., Orozco, R., Perniciano, M., Jurado, M., Melis, G. and Alcazar, J. *Accuracy of transvaginal ultrasound for diagnosis of deep endometriosis in the rectosigmoid: systematic review and meta-analysis.* Ultrasound in Obstetrics & Gynecology, vol. 47, no. 3, pages 281–289, 2016. (Cited in page 17.)

[Guo 2009] Guo, S.-W. *Recurrence of endometriosis and its control.* Human reproduction update, vol. 15, no. 4, 2009. (Cited in page 6.)

[Gupta 2005] Gupta, S., Chauhan, R. C. and Saxena, S. C. *Locally adaptive wavelet domain Bayesian processor for denoising medical ultrasound images using speckle modelling based on Rayleigh distribution.* Proc. Vis. Image Signal Process., vol. 152, no. 1, pages 129–135, 2005. (Cited in pages 60 and 69.)

[Gynaescope 2015] Gynaescope, W. *Laparoscopic Surgery for Endometriosis*, 2015. `https://www.wagynaescope.com.au/laparoscopic-surgery-for-endometriosis`. (Cited in pages ix and 7.)

[Hansen 2014] Hansen, M., Seshamani, S., Kroenke, C., Cheng, X., Chapman, T., Wilm, J., Rousseau, F. and Studholme, C. *A Unified Approach to Diffusion Direction Sensitive Slice Registration and 3-D DTI Reconstruction From Moving Fetal Brain Anatomy.* I E E E Transactions on Medical Imaging, vol. 33, no. 2, pages 272–289, 2014. [Online]. Available: http://dx.doi.org/10.1109/TMI.2013.2284014. (Cited in page 35.)

[Harada 2013] Harada, T. *Dysmenorrhea and endometriosis in young women.* Yonago acta medica, vol. 56, no. 4, page 81, 2013. (Cited in page 2.)

[Haribabu 2023] Haribabu, M., Guruviah, V. and Yogarajah, P. *Recent advancements in multimodal medical image fusion techniques for better diagnosis: an overview.* Current Medical Imaging, vol. 19, no. 7, pages 673–694, 2023. (Cited in page 60.)

[Haskins 2020] Haskins, G., Kruger, U. and Yan, P. *Deep learning in medical image registration: a survey.* Machine Vision and Applications, vol. 31, no. 1–2, January 2020. [Online]. Available: `http://dx.doi.org/10.1007/s00138-020-01060-x`. (Cited in page 30.)

[He 2013] He, K., Sun, J. and Tang, X. *Guided Image Filtering.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 6, pages 1397–1409, 2013. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2012.213. (Cited in pages 65 and 84.)

[Heinrich 2012] Heinrich, M. P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F. V., Brady, M. and Schnabel, J. A. *MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration.* Medical image analysis, vol. 16, no. 7, pages 1423–1435, 2012. (Cited in page 36.)

[Heinrich 2013] Heinrich, M. P., Jenkinson, M., Papież, B. W., Brady, S. M. and Schnabel, J. A. *Towards realtime multimodal fusion for image-guided interventions using self-similarities.* In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013: 16th International Conference, Nagoya, Japan, September 22-26, 2013, Proceedings, Part I 16, pages 187–194. Springer, 2013. (Cited in page 36.)

[Heldmann 2009] Heldmann, S. and Papenberg, N. *A Variational Approach for Volume-to-Slice Registration.* In Scale Space and Variational Methods in Computer Vision, pages 624–635, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. (Cited in page 28.)

[Huang 2011] Huang, P., Budd, C. and Hilton, A. *Global temporal registration of multiple non-rigid surface sequences.* In CVPR 2011, pages 3473–3480, 2011. (Cited in page 31.)

[James 2014] James, A. P. and Dasarathy, B. V. *Medical image fusion: A survey of the state of the art.* Information Fusion, vol. 19, pages 4–19, 2014. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1566253513001450`. Special Issue on Information Fusion in Medical Image Computing and Systems. (Cited in pages 60 and 65.)

[Joshi 2016] Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M. R., Kuemmerle, T., Meyfroidt, P., Mitchard, E. T.*et al. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring.* Remote Sensing, vol. 8, no. 1, page 70, 2016. (Cited in page 60.)

[Kadoury 2012] Kadoury, S., Zagorchev, L., Wood, B. J., Venkatesan, A., Weese, J., Jago, J. and Kruecker, J. *A model-based registration approach of preoperative MRI with 3D ultrasound of the liver for Interventional guidance procedures.* In 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), pages 952–955, 2012. (Cited in pages 30 and 36.)

[Kang 2014] Kang, S. H., Shafei, B. and Steidl, G. *Supervised and Transductive Multi-Class Segmentation Using p-Laplacians and RKHS Methods.* Journal of Visual Communication and Image Representation, vol. 25, no. 5, pages 1136–1148, 2014. (Cited in page 71.)

[Kaur 2021] Kaur, H., Koundal, D. and Kadyan, V. *Image Fusion Techniques: A Survey.* Archives of Computational Methods in Engineering, vol. 28, 01 2021.

[Online]. Available: http://dx.doi.org/10.1007/s11831-021-09540-7. (Cited in page 61.)

[Kaur 2022] Kaur, P., Panwar, G., Uppal, N., Singh, P., Shivahare, B. D. and Diwakar, M. *A Review on Multi-Focus Image Fusion Techniques in Surveillance Applications for Image Quality Enhancement.* In 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), pages 7–11, 2022. (Cited in page 60.)

[Khan 2020] Khan, S. A., Khan, M. A., Song, O.-Y. and Nazir, M. *Medical imaging fusion techniques: a survey benchmark analysis, open challenges and recommendations.* Journal of Medical Imaging and Health Informatics, vol. 10, no. 11, pages 2523–2531, 2020. (Cited in page 61.)

[Kinley 2015] Kinley, J. and Damadian, R. Gifted mind: the dr. raymond damadian story, inventor of the mri. New Leaf Publishing Group, 2015. (Cited in page 13.)

[Kiselev 2017] Kiselev, V. G. *Fundamentals of diffusion MRI physics.* NMR in Biomedicine, vol. 30, no. 3, page e3602, 2017. (Cited in page 16.)

[Klein 2007] Klein, S., Staring, M. and Pluim, J. P. W. *Evaluation of Optimization Methods for Nonrigid Medical Image Registration Using Mutual Information and B-Splines.* IEEE Transactions on Image Processing, vol. 16, no. 12, pages 2879–2890, 2007. [Online]. Available: http://dx.doi.org/10.1109/TIP.2007.909412. (Cited in page 33.)

[Komodakis 2007] Komodakis, N., Tziritas, G. and Paragios, N. *Fast, Approximately Optimal Solutions for Single and Dynamic MRFs.* 06 2007. (Cited in page 41.)

[Komodakis 2008] Komodakis, N., Tziritas, G. and Paragios, N. *Performance vs computational efficiency for optimizing single and dynamic MRFs: Setting the state of the art with primal-dual strategies.* Computer Vision and Image Understanding, vol. 112, no. 1, pages 14–29, 2008. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1077314208000982`. Special Issue on Discrete Optimization in Computer Vision. (Cited in page 41.)

[Krempien 2003] Krempien, R. C., Daeuber, S., Hensley, F. W., Wannenmacher, M. and Harms, W. *Image fusion of CT and MRI data enables improved target volume definition in 3D-brachytherapy treatment planning.* Brachytherapy, vol. 2, no. 3, pages 164–171, 2003. (Cited in page 60.)

[Larrea 2024] Larrea, E. P., Gómez, M. B., Salgado, S. and Goosman, C. A. *Diagnosis of endometriosis: analytical and imaging tests*, 2024. `https://www.invitra.com/en/endometriosis-diagnosis/`. (Cited in pages ix and 6.)

[Li 2013] Li, S., Kang, X. and Hu, J. *Image Fusion With Guided Filtering.* IEEE Transactions on Image Processing, vol. 22, no. 7, pages 2864–2875, 2013. [Online]. Available: http://dx.doi.org/10.1109/TIP.2013.2244222. (Cited in pages 66, 83, 84, 85, and 91.)

[Li 2017] Li, S., Kang, X., Fang, L., Hu, J. and Yin, H. *Pixel-level image fusion: A survey of the state of the art.* Information Fusion, vol. 33, pages 100–112, 2017. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1566253516300458`. (Cited in page 61.)

[Li 2018] Li, W., Du, J., Zhao, Z. and Long, J. *Fusion of Medical Sensors Using Adaptive Cloud Model in Local Laplacian Pyramid Domain.* IEEE transactions on bio-medical engineering, vol. PP, 09 2018. [Online]. Available: http://dx.doi.org/10.1109/TBME.2018.2869432. (Cited in page 64.)

[Li 2021] Li, Y., Zhao, J., Lv, Z. and Li, J. *Medical image fusion method by deep learning.* International Journal of Cognitive Computing in Engineering, vol. 2, pages 21–29, 2021. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S2666307420300280`. (Cited in page 64.)

[Liao 2009] Liao, M., Zhang, Q., Wang, H., Yang, R. and Gong, M. *Modeling deformable objects from a single depth camera.* In 2009 IEEE 12th International Conference on Computer Vision, pages 167–174, 2009. (Cited in page 31.)

[Liao 2012] Liao, C.-T., Fan, K.-H., Lin, C.-Y., Wang, H.-M., Huang, S.-F., Chen, I.-H., Kang, C.-J., Ng, S.-H., Hsueh, C., Lee, L.-Y.*et al. Impact of a second FDG PET scan before adjuvant therapy for the early detection of residual/relapsing tumours in high-risk patients with oral cavity cancer and pathological extracapsular spread.* European journal of nuclear medicine and molecular imaging, vol. 39, pages 944–955, 2012. (Cited in page 60.)

[Liu 2017] Liu, Z., Chai, Y., Yin, H., Zhou, J. and Zhu, Z. *A novel multi-focus image fusion approach based on image decomposition.* Information Fusion, vol. 35, pages 102–116, 2017. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1566253516300781`. (Cited in page 61.)

[Lladó 2012] Lladó, X., Ganiler, O., Oliver, A., Martí, R., Freixenet, J., Valls, L., Vilanova, J. C., Ramió-Torrentà, L. and Rovira, À. *Automated detection of multiple sclerosis lesions in serial brain MRI.* Neuroradiology, vol. 54, pages 787–807, 2012. (Cited in page 60.)

[Loza 2010] Loza, A., Bull, D., Canagarajah, N. and Achim, A. *Non-Gaussian model-based fusion of noisy images in the wavelet domain.* Computer Vision and Image Understanding, vol. 114, pages 54–65, 01 2010. [Online]. Available: http://dx.doi.org/10.1016/j.cviu.2009.09.002. (Cited in page 64.)

[M Thijssen 2014] M Thijssen, J. and L de Korte, C. *Cardiological ultrasound imaging.* Current Pharmaceutical Design, vol. 20, no. 39, pages 6150–6161, 2014. (Cited in page 8.)

[Ma 2016] Ma, J., Zhao, J. and Yuille, A. L. *Non-Rigid Point Set Registration by Preserving Global and Local Structures.* IEEE Trans. Image Process., vol. 25, no. 1, pages 53–64, January 2016. (Cited in page 71.)

[Ma 2018] Ma, K., Duanmu, Z., Yeganeh, H. and Wang, Z. *Multi-Exposure Image Fusion by Optimizing A Structural Similarity Index.* IEEE Transactions on Computational Imaging, vol. 4, no. 1, pages 60–72, 2018. [Online]. Available: http://dx.doi.org/10.1109/TCI.2017.2786138. (Cited in page 67.)

[Manjón 2010] Manjón, J. V., Coupé, P., A. Buades, A., Collins, D. L. and Robles, M. *MRI superresolution using self-similarity and image priors.* International journal of biomedical imaging, vol. 2010, 2010. (Cited in pages 59 and 69.)

[Marami 2014] Marami, B., Sirouspour, S., Fenster, A. and Capson, D. *Dynamic tracking of a deformable tissue based on 3D-2D MR-US image registration.* volume 9036, page 90360T, 03 2014. (Cited in page 36.)

[Maranur 2018] Maranur, J. R. *Image fusion using the discrete wavelet transform method.* International Journal of Scientific Development and Research (IJSDR), vol. 3, no. 9, 2018. (Cited in page 63.)

[Martin 2006] Martin, D. C. *Hysterectomy for treatment of pain associated with endometriosis.* Journal of minimally invasive gynecology, vol. 13, no. 6, pages 566–572, 2006. (Cited in page 5.)

[Mathiyalagan 2018] Mathiyalagan, P. and Suvitha, N. *Image Fusion Using Convolutional Neural Network with Bilateral Filtering.* In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pages 1–11, 2018. (Cited in page 65.)

[Mauri 2014] Mauri, G., Cova, L., Beni, S., Ierace, T., Tondolo, T., Cerri, A., Goldberg, S. and Solbiati, L. *Real-Time US-CT/MRI Image Fusion for Guidance of Thermal Ablation of Liver Tumors Undetectable with US: Results in 295 Cases.* Cardiovascular and interventional radiology, vol. 38, 05 2014. [Online]. Available: http://dx.doi.org/10.1007/s00270-014-0897-y. (Cited in page 104.)

[Meher 2019] Meher, B., Agrawal, S., Panda, R. and Abraham, A. *A survey on region based image fusion methods.* Information Fusion, vol. 48, pages 119–132, 2019. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1566253517307583`. (Cited in pages 60 and 65.)

[Meskine 2013] Meskine, F., Taleb, N., Miloud, c. e.-m., Kpalma, K. and Almhdie, A. *A Rigid Point Set Registration of Remote Sensing Images Based On Genetic Algorithms Hausdorff Distance.* 06 2013. (Cited in page 30.)

[Mishra 2015] Mishra, D. and Palkar, B. *Image Fusion Techniques: A Review.* International Journal of Computer Applications, vol. 130, pages 7–13, 11 2015. [Online]. Available: http://dx.doi.org/10.5120/ijca2015907084. (Cited in pages x and 62.)

[Mitra 2010] Mitra, J., Oliver, A., Martí, R., Lladó, X., Vilanova, J. C. and Meriaudeau, F. *Multimodal Prostate Registration Using Thin-Plate Splines from Automatic Correspondences.* In 2010 International Conference on Digital Image Computing: Techniques and Applications, pages 587–592, 2010. (Cited in page 32.)

[Moradi 2012] Moradi, M., Janoos, F., Fedorov, A., Risholm, P., Kapur, T., Wolfsberger, L. D., Nguyen, P. L., Tempany, C. M. and Wells, W. M. *Two solutions for registration of ultrasound to MRI for image-guided prostate interventions.* In 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pages 1129–1132, 2012. (Cited in page 29.)

[Movahed 2020] Movahed, R. A., Rezaeian, M., Javadifar, S. and Alimoradijazi, M. *A Face Recognition Framework Based on the Integration of Eigenfaces Algorithm and Image Registration Technique.* In 2020 27th National and 5th International Iranian Conference on Biomedical Engineering (ICBME), pages 26–30, 2020. (Cited in page 26.)

[Muzammil 2020] Muzammil, S. R., Maqsood, S., Haider, S. and Damaševičius, R. *CSID: A novel multimodal image fusion algorithm for enhanced clinical diagnosis.* Diagnostics, vol. 10, no. 11, page 904, 2020. (Cited in page 60.)

[Ni 2016] Ni, W. and Gao, X. *Despeckling of SAR Image Using Generalized Guided Filter With Bayesian Nonlocal Means.* IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 1, pages 567–579, 2016. [Online]. Available: http://dx.doi.org/10.1109/TGRS.2015.2462120. (Cited in pages 86, 87, and 89.)

[Osechinskiy 2011] Osechinskiy, S. and Kruggel, F. *Deformable registration of histological sections to brain MR images using a hybrid boundary-based slice-to-volume approach.* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference, vol. 2011, pages 4876–9, 08 2011. [Online]. Available: http://dx.doi.org/10.1109/IEMBS.2011.6091208. (Cited in page 32.)

[Paramanandham 2018] Paramanandham, N. and Rajendiran, K. *Multi sensor image fusion for surveillance applications using hybrid image fusion algorithm.* Multimedia Tools and Applications, vol. 77, 05 2018. [Online]. Available: http://dx.doi.org/10.1007/s11042-017-4895-3. (Cited in page 66.)

[Parazzini 1995] Parazzini, F., Ferraroni, M., Fedele, L., Bocciolone, L., Rubessa, S. and Riccardi, A. *Pelvic endometriosis: Reproductive and menstrual risk factors at different stages in Lombardy, northern Italy.* Journal of epidemiology and community health, vol. 49, pages 61–4, 03 1995. [Online]. Available: http://dx.doi.org/10.1136/jech.49.1.61. (Cited in page 5.)

[Patil 2011] Patil, U. and Mudengudi, U. *Image fusion using hierarchical PCA.* In 2011 International Conference on Image Information Processing, pages 1–6, 2011. (Cited in page 65.)

[Pereyra 2016] Pereyra, M., Schniter, P., Chouzenoux, E., Pesquet, J.-C., Tourneret, J.-Y., Hero, A. O. and McLaughlin, S. *A Survey of Stochastic Simulation and Optimization Methods in Signal Processing.* IEEE Journal of Selected Topics in Signal Processing, vol. 10, no. 2, page 224–241, March 2016. [Online]. Available: `http://dx.doi.org/10.1109/JSTSP.2015.2496908`. (Cited in page 102.)

[Porchetto 2017] Porchetto, R., Stramana, F., Paragios, N. and Ferrante, E. *Rigid Slice-To-Volume Medical Image Registration Through Markov Random Fields.* pages 172–185, 07 2017. (Cited in page 37.)

[Pratt 2018] Pratt, P., Ives, M., Lawton, G., Simmons, J., Radev, N., Spyropoulou, L. and Amiras, D. *Through the HoloLens™ looking glass: augmented reality for extremity reconstruction surgery using 3D vascular models with perforating vessels.* European Radiology Experimental, vol. 2, 12 2018. [Online]. Available: http://dx.doi.org/10.1186/s41747-017-0033-2. (Cited in pages xi and 103.)

[Qian 2019] Qian, L., Wu, J. Y., DiMaio, S. P., Navab, N. and Kazanzides, P. *A review of augmented reality in robotic-assisted surgery.* IEEE Transactions on Medical Robotics and Bionics, vol. 2, no. 1, pages 1–16, 2019. (Cited in page 103.)

[Qiu 2017] Qiu, C., Wang, Y., Zhang, H. and Xia, S. *Image Fusion of CT and MR with Sparse Representation in NSST Domain.* Computational and Mathematical Methods in Medicine, vol. 2017, pages 1–13, 11 2017. [Online]. Available: http://dx.doi.org/10.1155/2017/9308745. (Cited in page 67.)

[Ramin-Wright 2018] Ramin-Wright, A., Schwartz, A. S. K., Geraedts, K., Rauchfuss, M., Wölfler, M. M., Haeberlin, F., von Orelli, S., Eberhard, M., Imthurn, B., Imesch, P.*et al. Fatigue–a symptom in endometriosis.* Human reproduction, vol. 33, no. 8, pages 1459–1465, 2018. (Cited in page 2.)

[Reddy 2014] Reddy, S. N. and Chandrasekhar, C. *Design and Implementation of Image Fusion Technique Using DWT for Micro Air Vehicle.* International Journal of electronics & communication technology, vol. 5, march 2014. (Cited in page 63.)

[Rivaz 2014] Rivaz, H., Karimaghaloo, Z. and Collins, D. L. *Self-similarity weighted mutual information: a new nonrigid image registration metric.* Medical image analysis, vol. 18, no. 2, pages 343–358, 2014. (Cited in page 37.)

[Rivaz 2015] Rivaz, H., Chen, S. J.-S. and Collins, D. L. *Automatic Deformable MR-Ultrasound Registration for Image-Guided Neurosurgery.* IEEE Transactions on Medical Imaging, vol. 34, no. 2, pages 366–380, 2015. [Online]. Available: http://dx.doi.org/10.1109/TMI.2014.2354352. (Cited in pages 26 and 29.)

[Roche 2001] Roche, A., X.Pennec, Malandain, G. and Ayache, N. *Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information.* IEEE Trans. Med. Imag., vol. 20, no. 10, pages 1038–1049, 2001. [Online]. Available: http://dx.doi.org/10.1109/42.959301. (Cited in page 69.)

[Rueckert 1999] Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M. and Hawkes, D. *Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images.* Medical Imaging, IEEE Transactions on, vol. 18, pages 712 – 721, 09 1999. [Online]. Available: http://dx.doi.org/10.1109/42.796284. (Cited in pages 33 and 34.)

[Rueckert 2010] Rueckert, D. and Aljabar, P. *Nonrigid Registration of Medical Images: Theory, Methods, and Applications [Applications Corner].* IEEE Signal Processing Magazine, vol. 27, no. 4, pages 113–119, 2010. [Online]. Available: http://dx.doi.org/10.1109/MSP.2010.936850. (Cited in page 28.)

[Sabre 2018] Sabre, R. and Wahyuni, I. *Multi-focus image fusion using Laplacian Pyramid technique based on Alpha-Stable filter.* 01 2018. (Cited in pages x and 64.)

[Sahu 2014] Sahu, A., Bhateja, V., Krishn, A. and Himanshi. *Medical image fusion with Laplacian Pyramids.* In 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom), pages 448–453, 2014. (Cited in page 64.)

[Saleem 2012] Saleem, A., Beghdadi, A. and Boashash, B. *Image fusion-based contrast enhancement.* EURASIP Journal on Image and Video Processing, vol. 2012, 05 2012. [Online]. Available: http://dx.doi.org/10.1186/1687-5281-2012-10. (Cited in page 64.)

[Sanches 2012] Sanches, J. M., Laine, A. F. and Suri, J. S. Ultrasound imaging. Springer, 2012. (Cited in page 12.)

[Sanli 2017] Sanli, F. B., Abdikan, S., Esetlili, M. T. and Sunar, F. *Evaluation of image fusion methods using PALSAR, RADARSAT-1 and SPOT images for land use/land cover classification.* Journal of the Indian Society of Remote Sensing, vol. 45, pages 591–601, 2017. (Cited in page 60.)

[Sarvaiya 2009] Sarvaiya, J., Patnaik, S. and Bombaywala, S. *Image Registration by Template Matching Using Normalized Cross-Correlation.* In 2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies, pages 819–822, 2009. (Cited in page 28.)

[Scardapane 2014] Scardapane, A., Lorusso, F., Scioscia, M., Ferrante, A., Ianora, A. and Angelelli, G. *Standard high-resolution pelvic MRI vs. low-resolution pelvic MRI in the evaluation of deep infiltrating endometriosis.* European radiology, vol. 24, 07 2014. [Online]. Available: http://dx.doi.org/10.1007/s00330-014-3297-4. (Cited in pages ix and 18.)

[Sengupta 2022] Sengupta, D., Gupta, P. and Biswas, A. *A survey on mutual information based medical image registration algorithms.* Neurocomputing, vol. 486, pages 174–188, 2022. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0925231221016970`. (Cited in page 28.)

[Shahzad 2023] Shahzad, H., Bhatti, N., Phillips, F. and Khan, S. *Applications of Augmented Reality in Orthopaedic Spine Surgery.* Journal of the American Academy of Orthopaedic Surgeons, vol. Publish Ahead of Print, 04 2023. [Online]. Available: http://dx.doi.org/10.5435/JAAOS-D-23-00023. (Cited in page 103.)

[Singh 2008] Singh, A., Kruecker, J., Xu, S., Glossop, N., Guion, P., Ullman, K., Choyke, P. and Wood, B. *Initial clinical experience with real-time transrectal ultrasonography-magnetic resonance imaging fusion-guided prostate biopsy.* BJU international, vol. 101, pages 841–5, 05 2008. [Online]. Available: http://dx.doi.org/10.1111/j.1464-410X.2007.07348.x. (Cited in page 29.)

[Singla 2022] Singla, R. K., Kadatz, M., Rohling, R. and Nguan, C. *Kidney ultrasound for nephrologists: a review.* Kidney Medicine, vol. 4, no. 6, page 100464, 2022. (Cited in page 8.)

[Sinko 2018] Sinko, M., Kamencay, P., Hudec, R. and Benco, M. *3D registration of the point cloud data using ICP algorithm in medical image analysis.* In 2018 ELEKTRO, pages 1–6, 2018. (Cited in page 29.)

[Sivaramakrishna 2005] Sivaramakrishna, R. *3D Breast Image Registration — A Review.* Technology in cancer research & treatment, vol. 4, pages 39–48, 02 2005. [Online]. Available: http://dx.doi.org/10.1177/153303460500400106. (Cited in page 104.)

[Smith 1991] Smith, P., Bakos, O., Heimer, G. and Ulmsten, U. *Transvaginal ultrasound for identifying endometrial abnormality.* Acta obstetricia et gynecologica Scandinavica, vol. 70, no. 7-8, pages 591–594, 1991. (Cited in page 4.)

[Smolikova-Wachowiak 2005] Smolikova-Wachowiak, R., Wachowiak, M., Fenster, A. and Drangova, M. *Registration of two-dimensional cardiac images to pre-procedural three-dimensional images for interventional applications.* Journal of magnetic resonance imaging : JMRI, vol. 22, pages 219–28, 08 2005. [Online]. Available: http://dx.doi.org/10.1002/jmri.20364. (Cited in page 35.)

[So 2011] So, R. W., Tang, T. W. and Chung, A. C. *Non-rigid image registration of brain magnetic resonance images using graph-cuts.* Pattern Recognition, vol. 44, no. 10, pages 2450–2467, 2011. (Cited in page 35.)

[Spinczyk 2020] Spinczyk, D., Stronczek, M., Badura, A., Sperka, P., Krywalska, D., Wolinska, A., Krasoń, A., Fabian, S., Bas, M., Woloshuk, A.*et al. Initial clinical evaluation of image fusion based on rigid registration and supporting percutaneous liver tumor ablation.* Biocybernetics and Biomedical Engineering, vol. 40, no. 4, pages 1378–1390, 2020. (Cited in page 104.)

[Su 2013] Su, P., Yang, J., Lu, K., Yu, N., Wong, S. T. and Xue, Z. *A Fast CT and CT-Fluoroscopy Registration Algorithm With Respiratory Motion Compensation for Image-Guided Lung Intervention.* IEEE Transactions on Biomedical Engineering, vol. 60, no. 7, pages 2034–2041, 2013. [Online]. Available: http://dx.doi.org/10.1109/TBME.2013.2245895. (Cited in page 29.)

[Tan 2020] Tan, W., Tiwari, P., Pandey, H. M., Moreira, C. and Jaiswal, A. K. *Multimodal medical image fusion algorithm in the era of big data.* Neural computing and applications, pages 1–21, 2020. (Cited in page 61.)

[Tang 2018] Tang, X., Cai, F., Ding, D.-X., Zhang, L.-L., Cai, X.-Y. and Fang, Q. *Magnetic resonance imaging relaxation time in Alzheimer's disease.* Brain research bulletin, vol. 140, pages 176–189, 2018. (Cited in page 13.)

[Tao 2011] Tao, L. and Qian, Z.-Y. *An improved medical image fusion algorithm based on wavelet transform.* In 2011 Seventh International Conference on Natural Computation, volume 1, pages 76–78, 2011. (Cited in page 63.)

[Tavard 2014] Tavard, F., Simon, A., Leclercq, C., Donal, E., Hernández, A. I. and Garreau, M. *Multimodal registration and data fusion for cardiac resynchronization therapy optimization.* IEEE transactions on medical imaging, vol. 33, no. 6, pages 1363–1372, 2014. (Cited in page 104.)

[Tlig 2022] Tlig, M., Bouchouicha, M., Sayadi, M. and Moreau, E. *Infrared-visible images' fusion techniques for forest fire monitoring.* In 2022 6th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), pages 1–6, 2022. (Cited in page 60.)

[Tosti 2017] Tosti, C., Biscione, A., Morgante, G., Bifulco, G., Luisi, S. and Petraglia, F. *Hormonal therapy for endometriosis: from molecular research to bedside.* European Journal of Obstetrics & Gynecology and Reproductive Biology, vol. 209, pages 61–66, 2017. (Cited in page 5.)

[TRIVEDI 2023] TRIVEDI, G. and SANGHAVI, R. *Fusesharp: A multi-image focus fusion method using discrete wavelet transform and unsharp masking.* Journal of applied mathematics & informatics, vol. 41, no. 5, pages 1115–1128, 2023. (Cited in pages x and 63.)

[Uccella 2016] Uccella, S., Marconi, N., Casarin, J., Ceccaroni, M., Boni, L., Sturla, D., Serati, M., Carollo, S., Podesta'Alluvion, C. and Ghezzi, F. *Impact of endometriosis on surgical outcomes and complications of total laparoscopic hysterectomy.* Archives of gynecology and obstetrics, vol. 294, pages 771–778, 2016. (Cited in page 5.)

[Ultrasonography 2016] Ultrasonography. *Basics of Ultrasound*, 2016. `https://radiologykey.com/basics-of-ultrasound/`. (Cited in pages ix, 9, and 10.)

[Van den Elsen 1995] Van den Elsen, P., Maintz, J., Pol, E.-J. and Viergever, M. *Automatic registration of CT and MR brain images using correlation of geometrical features.* IEEE Transactions on Medical Imaging, vol. 14, no. 2, pages 384–396, 1995. [Online]. Available: http://dx.doi.org/10.1109/42.387719. (Cited in page 28.)

[Vercellini 2016] Vercellini, P., Buggio, L., Berlanda, N., Barbara, G., Somigliana, E. and Bosari, S. *Estrogen-progestins and progestins for the management of endometriosis.* Fertility and Sterility, vol. 106, no. 7, pages 1552–1571, 2016. (Cited in page 5.)

[Vidal 2019] Vidal, F., Mansouri, O. E., Kouamé, D. and Basarab, A. *On the Design of a Pelvic Phantom for Magentic Resononance and Utltrasound Image Fusion.* In 2019 IEEE International Ultrasonics Symposium (IUS), pages 2400–2403, 2019. (Cited in pages 43, 76, and 90.)

[Wang 2014] Wang, L., Li, B. and Tian, L.-f. *Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients.* Information Fusion, vol. 19, page 20–28, 09 2014. [Online]. Available: http://dx.doi.org/10.1016/j.inffus.2012.03.002. (Cited in page 68.)

[Watanabe 2017] Watanabe, H., Terada, T. and Tsukamoto, M. *Gesture Recognition Method Utilizing Ultrasonic Active Acoustic Sensing.* Journal of Information Processing, vol. 25, pages 331–340, 04 2017. [Online]. Available: http://dx.doi.org/10.2197/ipsjjip.25.331. (Cited in pages ix and 8.)

[Weinfeld 1998] Weinfeld, R., Johnson, S., Lucas, C. and Saksouk, F. *CT diagnosis of perihepatic endometriosis complicated by malignant transformation.* Abdominal imaging, vol. 23, pages 183–184, 1998. (Cited in page 4.)

[Wells 1996] Wells, W. M., Viola, P., Atsumi, H., Nakajima, S. and Kikinis, R. *Multi-modal volume registration by maximization of mutual information.* Medical Image Analysis, vol. 1, no. 1, pages 35–51, 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841501800049. (Cited in page 28.)

[Yamazaki 2013] Yamazaki, S., Kagami, S. and Mochimaru, M. *Non-rigid Shape Registration Using Similarity-Invariant Differential Coordinates.* In 2013 International Conference on 3D Vision - 3DV 2013, pages 191–198, 2013. (Cited in page 31.)

[Yang 2010] Yang, Q., Wang, L. and Ahuja, N. *A constant-space belief propagation algorithm for stereo matching.* In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 1458–1465, 2010. (Cited in page 35.)

[Yang 2012] Yang, B., Kim, M. and Madden, M. *Assessing optimal image fusion methods for very high spatial resolution satellite images to support coastal monitoring.* GIScience & Remote Sensing, vol. 49, no. 5, pages 687–710, 2012. (Cited in page 60.)

[Yang 2019] Yang, Y., Wu, J., Huang, S., Fang, Y., Lin, P. and Que, Y. *Multimodal Medical Image Fusion Based on Fuzzy Discrimination With Structural Patch Decomposition.* IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 4, pages 1647–1660, 2019. [Online]. Available: http://dx.doi.org/10.1109/JBHI.2018.2869096. (Cited in page 67.)

[Yavariabdi 2013] Yavariabdi, A., Samir, C., Bartoli, A., Da Ines, D. and Bourdel, N. *Mapping endometrial implants by 2D/2D registration of TVUS to MR images from point correspondences.* pages 576–579, 04 2013. (Cited in page 29.)

[Young 1984] Young, I. *Nuclear magnetic resonance imaging.* Electronics and Power, vol. 30, no. 3, pages 205–210, 1984. (Cited in page 13.)

[Zanardi 2003] Zanardi, R., Del Frate, C., Zuiani, C. and Bazzocchi, M. *Staging of pelvic endometriosis based on MRI findings versus laparoscopic classification according to the American Fertility Society.* Abdominal imaging, vol. 28, pages 733–742, 2003. (Cited in page 4.)

[Zhang 2021] Zhang, H., Xu, H., Tian, X., Jiang, J. and Ma, J. *Image fusion meets deep learning: A survey and perspective.* Information Fusion, vol. 76, pages 323–336, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253521001342. (Cited in page 64.)

[Zhao 2017] Zhao, X. *Image Fusion Based on IHS Transform and Principal Component Analysis (PCA) Transform.* In 2017 International Conference on Computer Technology, Electronics and Communication (ICCTEC), pages 304–307, 2017. (Cited in page 65.)

[Zhao 2023] Zhao, X., Jin, S., Bian, G., Cui, Y., Wang, J. and Zhou, B. *A Curvelet-Transform-Based Image Fusion Method Incorporating Side-Scan Sonar Image Features.* Journal of Marine Science and Engineering, vol. 11, page 1291, 06 2023. [Online]. Available: http://dx.doi.org/10.3390/jmse11071291. (Cited in page 64.)

[Zhong 2024] Zhong, Y., Zhang, S., Liu, Z., Zhang, X., Mo, Z., Zhang, Y., Hu, H., Chen, W. and Qi, L. *Unsupervised Fusion of Misaligned PAT and MRI Images via Mutually Reinforcing Cross-Modality Image Generation and Registration.* IEEE Transactions on Medical Imaging, vol. 43, no. 5, pages 1702–1714, 2024. [Online]. Available: http://dx.doi.org/10.1109/TMI.2023.3347511. (Cited in page 68.)

[Zhu 2022] Zhu, P., Ouyang, W., Guo, Y. and Zhou, X. *A Two-To-One Deep Learning General Framework for Image Fusion.* Frontiers in Bioengineering and Biotechnology, vol. 10, 2022. [Online]. Available: http://dx.doi.org/10.3389/fbioe.2022.923364. (Cited in page 65.)

[Zikic 2010] Zikic, D., Glocker, B., Kutter, O., Groher, M., Komodakis, N., Kamen, A., Paragios, N. and Navab, N. *Linear Intensity-based Image Registration by Markov Random Fields and Discrete Optimization.* Medical image analysis, vol. 14, pages 550–62, 08 2010. [Online]. Available: http://dx.doi.org/10.1016/j.media.2010.04.003. (Cited in pages 39 and 41.)

[Zondervan 2001] Zondervan, K. T., Cardon, L. R. and Kennedy, S. H. *The genetic basis of endometriosis.* Current Opinion in Obstetrics and Gynecology, vol. 13, no. 3, pages 309–314, 2001. (Cited in page 4.)

[Zou 2022] Zou, X., Feng, W., Quan, Y., Li, Q., Dauphin, G. and Xing, M. *A Multi-Level Synergistic Image Decomposition Algorithm for Remote Sensing Image Fusion.* In IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, pages 3754–3757, 2022. (Cited in page 60.)