Multiview Echography Restoration with Applications to 3D Breast and Cardiac Imaging

Pau Soler

▶ **To cite this version:**

Pau Soler. Multiview Echography Restoration with Applications to 3D Breast and Cardiac Imaging. domain_other. Télécom ParisTech, 2006. English. NNT : . pastel-00002462

HAL Id: pastel-00002462
https://pastel.hal.science/pastel-00002462

Submitted on 12 Jun 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PhD

Multiview Echography Restoration with Applications to 3D Breast and Cardiac Imaging

Restauration Echographique Multi-vues appliquée à l'Imagerie 3D du Sein et du Coeur

Pau Soler i Pla

PhD director : Isabelle Bloch

14th July 2006

Acknowledgements

This thesis is the result of the work of many people. It is my pleasure to show my gratitude to those who made it possible, especially to:

- Prof. Isabelle Bloch, thesis advisor, for sharing her geniality, expertise and kindness, for her extreme patience and availability, and for keeping me focused (or trying) when too many research ways were open;
- Dr. Olivier Gérard, co-director, for his enthusiasm, his incredible solution-finding abilities, and for inviting me to the world of 3D cardiac imaging;
- Dr. Eric Saloux, for his willingness to invent, share and improve, and his availability despite his busy-MD-agenda ;
- Alain Herment and Alison Noble, *rapporteurs*, for their careful revision of the manuscript, their improvements and their future research lines suggestions;
- Christian Barillot, *examinateur* and president of the jury, for his constructive criticism and his availability;
- David Cornwell, and to all the Marie Curie Fellowship program, for providing funds for this research;
- The whole Philips Medical Systems Research Paris team, leaded by Jean Pergrale, for making hard to forget those blackboard discussions, les *pots*, powerball and chess championships. Many thanks to Franck, Benoit, Nicolas, Roberto, Sherif, Stephane, Mathieu, Claire, Pierre, Raoul, Odile, Jean-Michel, Cécile, Françoise, Maxim, Laurance, Eric, Sham... And to Jean-Michel Lagrange, Claude Cohen-Bacrie and Rob Entrekin, who started this project, and to Pascal Allain, who significantly contributed to finish it. And of course, to Gaspar Delso, the Catalan companion, whose help could not be emphasized enough;
- All colleagues at ENST, for making my time at the university an unforgettable experience. To Elsa Angelini, for sharing her expertise, her vitality and her decisive contribution to the project. To Francis Schmidt, whose 'open-door' policy I have really appreciated. To all who built up a great team, including Oscar, Alejandro, Carlos, Penelope, Oliviers, Jamal, David, Penelope, Slim, Silvia, Cléo, Chloe, Tony, Judson, Thomas, Celina, Céline, Patricia... And very especially to Antonio Moreno, whose enormous help and endless discussions made possible to finish this thesis;

- All Parisian friends, who made my stay at the French capital simply wonderful. Gabriele, dónde está la chimba? Isra, Mirko, Elena, Lu, Pepe, Guido, Cristele, Sandra, Pedro, Cécile, Giordano, Mady, Hakim, Nadir, Benoit, Stephanie, Philippe, Emanuelle, Encarna, Isa, Jesus, Octavio... What a nice *souvenirs*! And especially to my favorite carioca. Fernanda, muito obrigado for teaching me the "art-of-the-finish";

- And family. So far, but so close. Bernat, Marta, Bruna, Toti, Tita, Martí, QK, Laia, Núria... you were there through all these times. Mamà, mother there is only one. Those priceless envelopes with "jamón serrano" really cheered up the 5 pl. Rungis. Papà, it is so sad you could not see the end of this thesis. *No els ho diguis, però crec que finalment he inventat la màquina de desmamar pollastres...*

To all, THANK YOU !!!!

Abstract

Ultrasound echography is one of the most widely used medical imaging modalities because it is non-invasive, real-time and cost-effective. However, it suffers from some image quality defects, such as speckle noise, limited spatial resolution and angle dependent tissue contrast. Spatial compounding, which consists in averaging acquisitions from different angles, has shown good performances in reducing speckle in 2D imaging. In our thesis, we extend spatial compounding capabilities to also improve spatial resolution and angle dependency, particularly for 3D breast and cardiac imaging.

We propose new techniques to combine acquisitions from different angles, following two approaches: multiview deconvolution and multiview fusion. Multiview deconvolution consists in solving the inverse problem of estimating the original volume from the different acquisitions, by modeling the degradation as a convolution of the original tissue with a space-varying point spread function (PSF). We propose a technique to estimate the PSF based on subspace techniques, adding *a priori* knowledge of the geometrical relationship between the different acquisitions and shape constraints. The inverse problem is regularized with an edge-preserving functional adapted to speckle noise statistics, and solved iteratively. The second approach, multiview fusion, consists in detecting the features of interest in each acquisition to build a combined volume. We propose different fusion techniques in spatial, spectral and wavelet coefficient domains.

We applied the developed techniques to 3D ultrasound breast imaging. Volumes were obtained by scanning the tissue with a linear array attached to a robotic platform. A well-known problem in breast imaging in the limited elevational resolution of linear arrays. We overcame this problem by scanning the tissue at different angles, and combining those acquisitions. A non-rigid registration process was developed to guarantee the alignment of the different acquisitions. Multiview deconvolution methods showed best results with respect to improvement of spatial resolution and signal-to-noise ratio (SNR). This leads to the improvement of important parameters for clinical practice, such as tissue delineation and contrast resolution both on phantom and *in vivo* data.

Finally, we applied the developed techniques to real-time three-dimensional (RT3D) ultrasound cardiac imaging. This imaging modality can be further improved by increasing the field-of-view (FOV) and heart wall contrast. With this aim, different views were acquired through different acoustic windows. In order to combine such acquisitions, a robust rigid registration algorithm was developed. Multiview deconvolution showed slightly better results than other techniques, improving heart wall contrast and SNR. This leads to improvement of clinically relevant parameters, such as tissue delineation and heart wall contrast.

Résumé

Nous présentons dans cette thèse de nouvelles techniques pour l'amélioration de la qualité d'images échographiques ultrasonores tridimensionnelles. L'échographie par ultrasons est l'une des technologie les plus utilisées dans l'imagerie médicale, car elle permet d'observer tous types de tissus mous. Elle repose sur la propagation d'ondes ultrasonores à travers le corps humain, qui sont réfléchies aux interfaces entre les tissus, créant des *échos* qui sont ensuite affichés pour leur interprétation médicale. Les avantages de cette technique sont nombreux. Entre autres, la technologie est :

- non-invasive : en plus de 50 ans de pratique, aucun effet co-latéral n'a été observé ;
- temps-réel : les images s'obtiennent de façon immédiate, ce qui permet d'interagir avec les organes et voir comment ils répondent. De plus, on peut ainsi observer les phénomènes rapides, comme le battement du cœur ou le flux sanguin ;
- relativement peu coûteuse par rapport à d'autres méthodes d'imagerie ;
- portable : les appareils peuvent être transportés, par exemple, dans les salles d'opérations ou au domicile des patients ;
- multi-organes : tous les types de tissus mous sont visualisables.

Cependant l'imagerie par ultrasons a quelques inconvénients. Du point de vue pratique, elle est dépendante du patient et du praticien. Par exemple, des facteurs comme l'obésité rendent difficile l'obtention d'échographies de bonne qualité. De plus, actuellement, le praticien tient la sonde échographique à la main, ce qui nécessite une certaine expérience pour obtenir des images de bonne qualité. Du point de vue de l'image, les limitations sont entre autres (voir section A.3) :

- La présence de *speckle* (terme anglais qu'on emploiera par la suite, se traduisant par *bruit de chatoiement* en français), qui est un bruit déterministe présent partout dans l'image. Il crée une texture même dans des surfaces homogènes et peut masquer des éléments importants pour le diagnostic. La taille du speckle dépend de plusieurs facteurs, mais notablement de la fréquence ultrasonore utilisée. La fréquence est limitée par l'atténuation qui détermine la profondeur à laquelle les images peuvent être prises à une certaine fréquence. Chez les humains cette profondeur peut atteindre une quinzaine de centimètres, ce qui limite la fréquence maximale utilisable et ainsi le niveau minimal du speckle.

- La résolution spatiale, qui fait référence à la plus petite taille d'éléments que l'on peut distinguer. Dans notre cas, la résolution dépend de la fréquence et de la taille de la sonde, et est donc ainsi limitée.

- La dépendance de l'angle. La visibilité des tissus dépend de l'orientation par rapport au faisceau d'ultrasons. S'ils sont orthogonaux, un fort écho sera produit dans la direction de la sonde, et le tissu sera très visible. S'ils sont plutôt parallèles, très peu d'énergie arrivera à la sonde et le tissu sera donc peu visible.

Des exemples de ces phénomènes peuvent être appréciés à la Figure 2.7.

Plusieurs techniques ont été proposées pour améliorer ces artefacts. L'objectif est d'obtenir une représentation fidèle de l'image originale du tissu à partir de l'image échographique obtenue. En particulier, la plupart des techniques sont dédiées à la réduction du speckle. Le filtrage spatial consiste à utiliser la statistique locale de l'image pour discriminer les réflecteurs réels des interférences créées par le speckle. Ce filtrage peut être de différents types : adaptif [64], à diffusion anisotropique [88, 110, 44] ou à seuillage de coefficients d'ondelettes [61, 15, 11].

Une autre approche consiste à utiliser plusieurs acquisitions avec une certaine diversité entre elles. Elle porte le nom de *compounding* (terme anglais qui peut se traduire comme *combinaison*). Elle peut prendre la forme de :

- combinaison temporelle (*temporal compounding*) [47], qui consiste à combiner différentes acquisitions faites à différents instants ;

- combinaison fréquentielle (*frequency compounding*) [40], qui consiste à utiliser différentes fréquences pour créer l'image échographique. Les différentes acquisitions ont une texture de speckle non corrélée, et leur moyenne réduit les effets de speckle ;

- combinaison spatiale (*spatial compounding*) [155, 116], qui consiste à obtenir des images selon différents points de vue, et qui produit aussi différentes textures non corrélées.

Nous allons nous intéresser particulièrement à cette dernière technique, le *spatial compounding*, qui a prouvé son efficacité en imagerie bidimensionnelle ces dernières années.

Notre contribution intervient principalement sur deux axes :

1. Méthodes. Développer de nouvelles techniques de *spatial compounding*, avec l'objectif d'améliorer le rapport signal-sur-bruit mais aussi la résolution spatiale. On propose deux approches différentes :

    - *Déconvolution multi-vues aveugle* (qu'on appellera dans le texte *blind multiview deconvolution*), consistant à résoudre le problème inverse, où l'on suppose que chaque acquisition est une version dégradée du tissu original, et on recherche l'image la plus vraisemblable.

- *Fusion multi-vues* (qu'on appellera dans le texte *Multiview fusion*), où l'on construira un volume à partir des caractéristiques présentes dans chaque acquisition.

2. Applications. Démontrer la valeur ajoutée de ces méthodes dans de nouvelles applications, en particulier sur :

   - Imagerie 3D du sein. Typiquement en imagerie 2D, l'obtention des volumes 3D du sein reste un défi. On utilisera une sonde linéaire collée à un bras robotique pour effectuer un scan de la totalité du tissu. Cependant la résolution est limitée dans le plan d'élévation (voir annexe A.3.1) et pose quelques problèmes, que l'on va résoudre avec les méthodes proposées.
   - Imagerie 3D+T du cœur. De nouvelles sondes temps-réel 3D pour l'imagerie du cœur ont été récemment introduites sur le marché. La combinaison des acquisitions à travers différentes fenêtres acoustiques (voir section 6.2.3) permet d'augmenter le champ de vue, le contraste des parois et le rapport signal-sur-bruit.

   En pratique, chaque application nécessite une étape préliminaire de recalage des différentes acquisitions avant la combinaison, c'est pourquoi des techniques spécifiques ont été développées.

Voici le contenu de chacun des chapitres.

Chapitre 1 : Introduction

Nous présentons ici la problématique générale, avec le même niveau de détail que dans ce résumé.

Chapitre 2 : Multiview Deconvolution

Nous introduisons la déconvolution multi-vues. Nous commençons par un rappel de la problématique classique de déconvolution, où l'on souhaite restaurer une image à partir d'une acquisition modélisée comme étant dégradée par une opération de convolution et l'addition de bruit. On présente différentes solutions proposées dans la littérature, et on conclut en remarquant que les fréquences qui ont été trop atténuées seront difficiles (voire impossible) à restaurer.

Ensuite, le cas multi-canaux est discuté. On dispose de différentes acquisitions de la même image originale mais convoluée avec des noyaux différents. Dans ce cas, le problème est moins mal posé, au sens où les informations manquantes d'un canal peuvent être complétées par d'autres canaux. Dans certaines conditions, il est même possible de restaurer la forme exacte de l'image originale. On présente aussi des techniques de régularisation permettant le contrôle de l'amplification du bruit, et ainsi que la préservation des bords.

On s'attache finalement au cas particulier des cas multi-canaux avec multi-vues, où l'on dispose de certaines relations spatiales entre les différents noyaux. Cette configuration, spécifique à nos applications, nous permet de mieux comprendre la contribution de chaque acquisition au plan spectral, et les contraintes géométriques pour capturer les caractéristiques qui nous intéressent.

Chapitre 3 : Blind Multiview Deconvolution

Afin d'exploiter les techniques décrites au chapitre précédent, la connaissance des différents noyaux, qu'on appellera PSFs (de l'anglais *Point Spread Function*, ou *tache focale* en français), est nécessaire. On se demande alors s'il est possible de les estimer à partir des seules acquisitions. Pour répondre à cette question, nous étudions la littérature de la déconvolution aveugle (*blind deconvolution*, en anglais), pour les cas mono et multi canaux. Nous nous intéressons en particulier aux techniques de sous-espace, qui nous donnent une solution fermée des PSFs uniquement à partir des acquisitions.

Malheureusement, les techniques de sous-espace sont assez sensibles au bruit et à l'estimation du support de la PSF. Pour améliorer leur performance, nous ajoutons les contraintes géométriques des différentes PSFs afin de réduire le nombre d'inconnus. Pour les rendre encore plus robustes nous introduisons des contraintes de forme, en particulier de forme gaussienne, qui nous permettront d'utiliser ces techniques dans les conditions réelles de nos applications.

Nous terminons ce chapitre avec l'évaluation des techniques de déconvolution multi-vues aveugle sur des images synthétiques créées avec simulation du bruit de speckle et dégradation spatiale des images d'échographie ultrasonore.

Chapitre 4 : Multiview Fusion

Dans ce chapitre nous abordons une autre approche de restauration: la fusion multi-vues. On crée alors un volume fusionné à partir des caractéristiques présentes dans chacune des acquisitions. On espère que ces techniques seront plus simples car il ne faut pas estimer l'ensemble des PSF. Nous disposons en effet de solutions directes et non itératives, plus rapides que la déconvolution multi-vues. Nous présentons plusieurs méthodes :

- Moyenne pondérée (*Weighted averaging* en anglais), consistant à affecter un poids différent à chaque vue en fonction d'une mesure de saillance et de speckle.
- Moyenne généralisée (*Generalized averaging* en anglais), qui permet une transition douce entre des opérateurs de minimum, moyenne et maximum en fonction de la différence entre les différentes acquisitions.
- Fréquences Maximales (*Maximum frequency* en anglais), qui préservent dans chaque acquisition les parties du spectre qui contiennent le plus d'énergie.
- Fusion par ondelettes (*Wavelet Fusion* en anglais), qui préserve les coefficients de la transformée par ondelettes contenant le plus d'énergie. Pour ne pas préserver

aussi tout le bruit, un seuillage des coefficients est effectué indépendamment à chaque niveau de la transformation.

On termine ce chapitre par une étude comparative de toutes ces méthodes, en incluant la déconvolution multi-vues aveugle. Cette dernière donne les meilleures performances en toutes conditions. De plus, elle est la seule qui arrive à égaliser la dégradation du système. Cependant, les techniques de fusion donnent des résultats acceptables avec un temps d'exécution plus faible.

Dans le chapitre suivant, nous appliquons ces techniques à de nouvelles applications.

Chapitre 5 : 3D Breast Ultrasound Imaging

L'imagerie du sein est d'extrême importance pour le dépistage du cancer. Dans ce chapitre, nous utilisons les techniques développées pour améliorer la qualité des images échographiques 3D du sein. L'obtention des volumes échographiques du sein présente plusieurs obstacles, les sondes actuelles étant trop petites par rapport à la taille des tissus à acquérir en image. Le système utilisé comporte une sonde linéaire fixée à un bras mécanique, balayant la sonde linéairement en créant un volume à partir de différentes coupes. En plus du speckle, les images ont une faible résolution sur le plan d'élévation, due à la petite ouverture selon cette dimension.

Pour résoudre ces problèmes nous avons proposé d'acquérir des images dans différentes orientations, et d'appliquer les techniques développées dans cette thèse. On a d'abord analysé les données obtenues pour vérifier les hypothèses posées lors de l'étude théorique. L'approximation gaussienne comme forme de PSF se révèle assez valable (voir section 5.4.1), et le bruit de speckle peut être modélisé comme un bruit additif gaussien coloré (voir section 5.4.2).

Avant de fusionner les différentes acquisitions, nous avons utilisé un algorithme de recalage élastique fondé sur les FFD (Free Form Deformations). A première vue, les effets de cette étape sont minimes car les images sont déjà initialement très bien recalées. Néanmoins des artefacts potentiels peuvent apparaître, dus aux mouvements involontaires du patient et qu'il faut corriger.

Une fois les volumes bien recalés, ils sont alors combinés avec les techniques présentées. Sur des images de fantômes, on a vérifié la diminution de la taille de la PSF (amélioration de la résolution spatiale) et la réduction du bruit de speckle. La technique de déconvolution multi-vues aveugle présente les meilleurs résultats. On a pu vérifier sur les images *in vivo* que cela se traduisait par une meilleure délimitation des tissus et une augmentation du contraste des masses, paramètres critiques pour le diagnostic oncologique.

Chapitre 6 : 3D Heart Ultrasound Imaging

Les maladies cardiovasculaires sont la première cause de mortalité dans le monde. Récemment, des systèmes 3D en temps-réel échocardiographiques capables d'obtenir des séquences de volumes et permettant une meilleure estimation des fonctions cardiaques ont été introduits. Les limitations de cette technologie sont : (i) un champ de vue ne contenant pas tout le cœur, en particulier avec les patients aux cœurs dilatés,

(ii), des ombres dues à l'impossibilité des ultrasons de traverser les côtes et l'air des poumons, ou dues à la dépendance à l'orientation des parois, et (iii) la présence de speckle.

Pour améliorer ces points, nous proposons de faire des acquisitions à travers différentes fenêtres acoustiques. Leur combinaison, de type mosaïque 4D, permet d'augmenter le champ de vue, et d'améliorer le contraste des parois et le rapport signal-sur-bruit.

Dans cette application, la partie de recalage est beaucoup plus difficile que pour l'application du sein, car les différentes acquisitions ne sont pas dans le même repère. On a développé une méthode de recalage rigide suffisamment robuste pour tenir compte des difficultés de cette application. Ensuite, les images sont fusionnées avec les techniques développées. Les résultats montrent que la déconvolution multi-canaux aveugle est également la plus performante sur cette application.

Chapitre 7 : Conclusions and Future Work

Finalement, nous résumons les conclusions et nous traçons les futures lignes de travail.

Conclusions :

- Nous avons prouvé que l'on pouvait à la fois réduire la puissance de speckle et améliorer la résolution spatiale à partir des acquisitions depuis différents points de vue.
- Nous avons présenté plusieurs façons de combiner ces acquisitions, et la deconvolution multi-vues aveugle s'est révélée être la plus performante avec des images synthétiques de fantôme et *in vivo*.
- Il est effectivement possible d'améliorer la résolution dans le plan d'élévation des sondes linéaires par restauration multi-vues.
- La restauration multi-vues des échocardiographies 3D+T permet d'augmenter le champ de vue et d'améliorer le contraste des parois et le rapport signal-sur-bruit.

Futures lignes de travail :

- Estimation automatique des hyperparamètres de régularisation.
- Autres schémas pour l'estimation de PSF et de l'image.
- Validation clinique sur des cas pathologiques pour l'imagerie du sein, et des nouvelles maladies pour l'imagerie du cœur.
- Application de ces techniques pour l'imagerie échographique 2D.
- Application de ces techniques dans d'autres domaines, comme l'obstétrique.

Table of Contents

CHAPTER 1

Introduction

In this thesis we present new techniques to improve the image quality of 3D ultrasound echography. This imaging technique is one of the most widely used since it permits to observe all kinds of soft tissues in a non-invasive way. It is based on the propagation of ultrasonic acoustic waves through the body, which are reflected by the tissue boundaries, creating *echoes* that are visually displayed for medical interpretation. In more than 50 years of usage in clinical practice, no long-term side effects have been reported. In comparison to other imaging techniques, it is relatively low cost, devices are portable and available almost everywhere.

However, ultrasound echography has several limitations. From an imaging point of view, exams are corrupted by a particular type of noise named speckle. This artifact causes a textured patterning that can hide important features for diagnosis. The amount of speckle is determined by the frequency of the ultrasound wave, which is constrained by the depth at which the tissue needs to be imaged. For applications on humans, the depth varies from 2 to 15 centimeters, bounding the frequency limits, thus the amount of speckle. The spatial resolution, which refers to the size of smallest element that can be distinguished, is also determined by the frequency and the size of the probe used, which again are subject to physical constraints. Also, the visibility of tissues depend on their orientation with respect to the ultrasound probe. Tissues perpendicular to the ultrasound beam create strong echoes, and are therefore very visible, and parallel ones create weak echoes, and are therefore poorly visible. Besides the imaging limitations, echography is patient and practitioner dependent. Factors such as obesity can greatly reduce the echogenicity of the patient, resulting into a difficult exam to interpret. Indeed, the practitioner needs to have advanced skills to properly place the probe on the body of the patient. Otherwise, relevant elements for clinical diagnosis may not be present in the echographic exam.

Many techniques have been proposed to overcome the aforementioned imaging limitations, including: spatial filtering, temporal integration, frequency compounding and spatial compounding. The goal is always to restore an image of original tissue, overcoming the limitations of the imaging system. Typically, the goal is focused on reducing speckle. Spatial filtering consists in exploiting the local statistics to discriminate homogeneous areas from real reflectors. This kind of filtering can take many forms, such as adaptive filters [64], anisotropic diffusion [88, 110, 44] or wavelet shrinkage [61, 11].

Applications that include a temporal dimension, averaging different acquisitions at different instants can improve the signal-to-noise ratio [47]. The frequency compounding technique [40] consists in imaging the tissue at different frequencies, obtaining uncorrelated speckle patterns in each acquisition. The same principle is performed in the spatial compounding technique [155, 116], where several acquisitions are obtained from different points of view. The tissue is either insonified from different parts of the probe or with different angles, obtaining uncorrelated speckle patterns. Averaging the different views increases the signal-to-noise ratio by $\sqrt{M}$ [136], where M is the number of views used. We will focus on the spatial compounding technique, which has proved to be very effective with two-dimensional imaging in the recent years, with the aim of further developing this technique.

Our contribution can be mainly divided into two axes: (i) develop new techniques of spatial compounding with the goal of improving the signal-to-noise ratio but also the spatial resolution and the field of view, and (ii) demonstrate their value in new applications, particularly on 3D breast and cardiac imaging.

Regarding the combination techniques (i), we follow two different approaches:

- We state the restoration problem as an inverse problem, where the different views are degraded versions of the original image, and we recover the most probable original image. This leads us to estimate the degradation of the system, under some assumptions on our data. This technique belongs to the field of *blind multiview deconvolution*.

- We propose to construct a combined volume from the features detected in each of the acquisitions. We propose techniques both in the spatial and transformed domains, denoting them *multiview fusion* techniques.

Regarding the applications (ii), we used the developed techniques for:

- 3D breast imaging. Volumetric imaging of the breast tissue is a challenging task, since there are no probes for this application. The breast tissue is scanned with a linear probe attached to a mechanical arm, but the limited resolution of linear arrays in the elevational plane is a critical issue. The developed techniques overcome this difficulty. Indeed, the combination of the different acquisition improves the signal-to-noise ratio.

- 3D+T cardiac imaging. We applied the developed techniques to the recently introduced 3D cardiac ultrasound, combining acquisitions from different acoustic windows. In this way, the field of view is improved, as well as the tissue delineation, the signal-to-noise ratio and the spatial resolution.

In practice, the different views need to be perfectly aligned before being combined. Otherwise, blurring artefacts may appear and reduce the benefits of the combination techniques. Therefore, registration issues and techniques will be described for each application.

This document is organized as follows.

In Chapter 2 we introduce the concept of multiview, as different acquisitions obtained with different angles. This diversity turns into a diversity in the frequency domain which allows the restoration at least at these parts of the spectrum. The intrinsic problems of single channel and multichannel deconvolution theory are reviewed, as well as regularization techniques. The obtained solutions have never been used previously for spatial compounding in ultrasound imaging.

In order to apply the techniques described in Chapter 2, we need to estimate the point spread function of the system. We review the blind multichannel techniques for this purpose in Chapter 3, highlighting their limitations. In particular, we introduce the multiview constraints into subspace techniques to improve their robustness to noise and support estimation. Without this specific particularization, multichannel techniques would not be feasible on real data.

In Chapter 4 we propose new alternative techniques to build a fused volume by detecting and choosing the features of interest of each view. We propose four new different techniques, namely weighted averaging, generalized averaging, maximum frequency and wavelet coefficient fusion. These techniques are simple, since they do not require an iterative schema to find the solution. Moreover, they do not depend on the estimation of the point spread function.

We proceed to apply the developed techniques to real applications. In Chapter 5 we obtain high resolution ultrasound volumetric images of the breast, including high-resolution C-scans. It would not be possible to obtain such images with a single scan. A non-linear registration scheme is implemented to register the different views. The limited resolution in the elevation dimension is improved by applying the new developed techniques, which also improve the signal-to-noise ratio.

In Chapter 6 we propose the multiview restoration techniques for cardiac applications. Acquisitions from different acoustic windows are obtained and combined with the proposed techniques, which permit to increase the field of view, the myocardium contrast and the signal-to-noise ratio. In order to combine such acquisitions, a robust rigid registration algorithm is developed.

Finally, conclusions and future research lines are provided in Chapter 7.

CHAPTER 2

Multiview Deconvolution

2.1 Introduction

Imaging systems are sensitive to certain physical properties of the imaged object, within certain limitations due to the nature of their sensors. Therefore, the output of the imaging systems provides partial information on the real object. Image restoration techniques try to recover the original image, taking out the system degradation. It is a key area in signal and image processing, and its applications include almost all imaging disciplines [14]. The degradation induced by the system may be complex and difficult to describe. However, under the hypotheses of spatial invariance and linearity, the observed image v_d can be modeled as a convolution operation between the original signal v with the system's point spread function (PSF) h [57]:

$$v_d = h * v \tag{2.1}$$

The formation of ultrasound images is a rather complicated process, as described in Appendix A.2. In this thesis we focus on the restoration of log-compressed envelope images. In this case, the spatial degradation of the system can be considered as a convolution with a spatially varying PSF [79]. In this chapter, however, we simplify the discussion to the case of spatially invariant PSF and tackle this problem in Chapters 5 and 6 for each specific application.

Deconvolution, a particular case of restoration, consists in inverting the convolution operation, that is, obtain v from v_d and h. It has been extensively reviewed in the literature [83], and successfully applied to a diversity of domains including astronomy [144], digital photography [75], microscopy [41], seismic sciences [71] and medical imaging [53]. Another particular case of restoration, related to but different from deconvolution, is image reconstruction. It may refer to combining several acquisitions of the same scene, such as in computer tomography imaging [135], where the image is reconstructed from its projections.

In order to make v_d as close as possible to v, our strategy is to acquire not one but M times with different orientations. These different acquisitions, or *views*, have complementary information that combined lead to a better image quality of the system.

We denote these acquisitions by v_i:

$$v_i = h_i * v, \quad i = 1..M \tag{2.2}$$

This second scheme receives the name of multichannel deconvolution, while the first is referred to as single channel deconvolution, or simply deconvolution. Note that in this case there are M channels providing information, and only one unknown v.

This chapter starts by briefly discussing the state-of-the-art of the single channel deconvolution case, to understand its intrinsic limitations. The multichannel case is presented in detail, with references to the single channel case to clarify concepts. The theory is developed generically, with some particularizations to echography imaging. Further on, we introduce the concept of multiview deconvolution, which refers to the fact that the different channels are in fact acquisitions from different points of view. The application to real cases is treated in detail in Chapters 5 and 6. Moreover, within this chapter, we assume that the spatial response h is always known, and we tackle the blind case -where h is unknown- in Chapter 3.

2.2 Single Channel Case

If the system degradation is assumed to be linear and shift invariant, the relationship between the original signal v and the degraded signal v_d can be described, in the continuous domain, as the convolution integral. For one dimensional signals, we have:

$$v_d(x) = h(x) * v(x) = \int_{-\infty}^{\infty} h(x-\tau)v(\tau)d\tau \tag{2.3}$$

In the discrete domain, the relationship is described by the convolution summation. Again, for one dimensional signals, we have:

$$v_d[x] = v[x] * h[x] = \sum_{\tau=-\infty}^{\infty} h[x-\tau]v[\tau] \tag{2.4}$$

If additive noise is supposed, the acquired signal can be modeled as:

$$v_{dn} = h * v + n \tag{2.5}$$

and its Fourier transform is

$$V_{dn} = H \cdot V + N \tag{2.6}$$

where H, the Fourier transform of the PSF h, is also known as optical transfer function (OTF). Typically H has the form of a low pass filter, describing different kinds of blur such as motion, out-of-focus or atmospheric turbulence. Figure 2.1 shows an example of an original and a degraded image. As it can be seen, since many features have been lost, it is an ambitious goal to recover the original image using only the observed image.

a) original image

b) blurred and noised image

Figure 2.1: The original image a) has been blurred with an horizontal Gaussian kernel and 10% noise added, resulting in b). The goal of deconvolution techniques is to invert this degradation and recover the original image a) from b).

Equation (2.3) is a Fredholm integral of the first kind [16], so recovering v from v_d and h is in general an ill-posed problem in the sense of Hadamard [63]. This means that the solution may not exist, may not be unique or may not be stable. For instance, if H is a perfect low-pass filter with cut-off frequency w_0, signals V' equal to V for $w < w_0$, but with arbitrary values for $w \geq w_0$, would also be solution of Equation (2.6). In other words, if some information is missing, many solutions can be found which are solution of Equation (2.6). Without *a priori* knowledge about the original signal, the later will not be recovered.

2.2.1 Deterministic approaches

Least Squares Solution

To find a solution of Equation (2.5), denoted as $\tilde{v}$, one may consider minimizing the differences between v_d and $h * \tilde{v}$. We are looking for a $\tilde{v}$ which, convolved with the system's PSF h, is most similar to the acquisition v_d. Using the $L2$-norm as distance between two signals results in:

$$\tilde{v} = \arg\min \|v_{dn} - h * \tilde{v}\|^2 \tag{2.7}$$

By Parseval's theorem, the system can also be expressed in the Fourier domain, as follows:

$$\tilde{V} = \arg\min \left\| V_{dn} - H \cdot \tilde{V} \right\|^2 \tag{2.8}$$

The solution to Equation (2.8) is [144]:

$$\tilde{V}_{LS}(\mathbf{w}) = \frac{H(\mathbf{w})^*}{|H(\mathbf{w})|^2} \cdot V_{dn}(\mathbf{w}) = \frac{H(\mathbf{w})^*}{|H(\mathbf{w})|^2} \cdot V(\mathbf{w}) + \frac{H(\mathbf{w})^*}{|H(\mathbf{w})|^2} \cdot N(\mathbf{w}) \tag{2.9}$$

This technique, also known as Fourier Quotient, illustrates the intrinsic problem of deconvolution: small values of $|H(\mathbf{w})|^2$ turn into noise amplification. Indeed, $\tilde{V}_{LS}(w)$ is not even defined for $|H(\mathbf{w})|^2 = 0$, the zeros of the OTF. Such problems make this technique unpractical. On the other hand, in the noiseless case and if $|H(\mathbf{w})|^2$ has no zeros in all the spectrum, deconvolution is exact. Unfortunately, this is a rare case in practice.

Constrained Least Squares Solution

To overcome the mentioned problems with the least squares solution, the solution space can be constrained by adding some *a priori* information about the original signal v. Typically, it is desired to limit the high frequency content of $\tilde{V}$ due to low values of $H(w)$. Given an operator C, also known as regularization function, we can express the constrained least squares problem as a minimization of Equation (2.7) subject to:

$$\left\| C \cdot \tilde{V} \right\|^2 \leq E \tag{2.10}$$

where E is the maximum amount of energy at the band-pass of C admitted in the solution $\tilde{V}$. The function C usually takes the form of a high-pass filter.

There exist many approaches to find a solution of the system composed by Equations (2.8) and (2.10). The Tikhonov-Miller method [151] consists in defining an energy J as a weighted sum of Equation (2.8), which is referred to the data fidelity term, and Equation (2.10), which is referred to as the regularization term.

$$J(\tilde{V}) = \left\| V_{dn} - H \cdot \tilde{V} \right\|^2 + \lambda \cdot \left\| C \cdot \tilde{V} \right\|^2 \tag{2.11}$$

The λ parameter determines the trade-off between these two terms. Along with the other parameters C may depend on, it is known as a regularization hyperparameter.

The solution that minimizes Equation (2.11) for linear operators is [144]:

$$\tilde{V}_{RLS}(w) = \arg\min J = \frac{H(w)^*}{|H(w)|^2 + \lambda |C(w)|^2} \cdot V_{dn}(w) \tag{2.12}$$

Again, $\tilde{V}_{RLS}$ is only defined if the denominator $|H|^2 + \lambda |C|^2 > 0, \ \forall w$, which is equivalent to the Hessian of J being definite positive. Notice, however, that in this case it suffices that $C(w)$ is not zero at the zeros of $H(w)$ for $\tilde{V}_{RLS}$ to be defined.

The solution of Equation (2.11) can also be obtained by iterative techniques. For convex energies, gradient descent techniques (e.g. conjugate gradient descent [125]) will converge to the global minimum in a fast way. If the energy is not convex, the convergence to a global minimum is difficult to be guaranteed analytically, and other techniques (e.g. graduated non convexity [27]) should be used.

Iterative solutions have some advantages such as not requiring the computation of big matrices and being numerically more stable. Additional constraints can be imposed to the intermediate solution, such as non-negativity. Indeed, the number of iterations can also be used as a regularization parameter.

2.2.2 Stochastic approaches

Up to now, the solution, signal and noise have been regarded as deterministic quantities. Considering signal and noise as stochastic processes, constraints can be specified using their statistical characteristics.

Wiener filter

Defining by $S_v(w)$ and $S_n(w)$ the spectral power of the signal process v and noise process n respectively, we can define C in Equation (2.10) as $C(w) = \sqrt{|S_n(w)| \,/\, |S_v(w)|}$. For $\lambda = 1$, the Wiener filter solution is obtained:

$$\tilde{V}_{Wiener} = \frac{H^*}{|H(w)|^2 + \frac{|S_n(w)|}{|S_v(w)|}} \cdot V_{dn}(w) = H_{Wiener}(w) \cdot V_{dn}(w) \tag{2.13}$$

Interpreting $C(w) = \sqrt{|S_n(w)| \,/\, \left|S_{v(w)}\right|}$ in the context of Equation (2.10), it can be seen that frequencies where noise is predominant are attenuated. Wiener solution $\tilde{\mathsf{V}}_{\mathsf{Wiener}}$ is the optimal solution to the least squares criterion 2.7 if the noise n follows a Gaussian zero-mean distribution.

Hyperparameter estimation

In the general case, the optimal choices for the hyperparameters and regularization functions in Equation (2.12) depend on the signal and noise. Hyperparameter estimation techniques [28] use signal and noise statistics, and thus, are considered as stochastic algorithms.

Bayesian Approach

Given v_d, the objective is to find the most probable v, that is, maximizing the probability $p(v|v_d)$. Defining by $p(v_d)$ the probability of the blurred image and by $p(v)$ the probability of the original image, the Bayes theorem [20] establishes the following relationship:

$$p(v|v_d) = \frac{p(v_d|v)p(v)}{p(v_d)} \tag{2.14}$$

The maximum likelihood (ML) solution is:

$$\tilde{v}_{ML} = \arg\max p(v_d|v) \tag{2.15}$$

The maximum *a posteriori* (MAP) solution is:

$$\tilde{v}_{MAP} = \arg\max p(v_d|v)p(v) \tag{2.16}$$

In both cases $p(v_d)$ is considered constant and ignored in the maximization process. The ML solution is the same as the MAP if $p(v)$ is assumed to be an uniform probability density function. Actually, $p(v)$ plays the role of regularization, expressing the knowledge about v. The regularization operators described in Section 2.2.1 can be expressed in the Bayesian framework as:

$$p(v) \ \alpha \ e^{-\frac{\lambda}{2}\|Cv\|^2} \tag{2.17}$$

In other words, the MAP solution is the regularized ML solution.

In the case where the additive noise is supposed to follow a Gaussian distribution with zero mean and variance σ_N^2, the *a posteriori* probability is

$$p(v_d|v) = \frac{1}{\sqrt{2\pi}\sigma_N} \ e^{-\frac{(v_{dn}-h*v)^2}{2\sigma_N^2}} \tag{2.18}$$

and the ML solution is equivalent to the least squares solution, since maximizing Equation (2.18) is equivalent to solving Equation (2.7).

If the original image v is also assumed to follow a Gaussian distribution with zero mean and variance σ_V^2, then the MAP solution is equivalent to the Wiener filter solution.

Similarly, the expectation maximization (EM) method is obtained by assuming a Poisson distribution for n, and maximizing $p(v_d|v)$ using the Picard iteration [73]. This method is also known as the Richardson-Lucy algorithm [135], and has the interesting property that, by construction, the solution is always positive.

2.2.3 Deconvolution of Ultrasound Images

The deconvolution of B-scan ultrasound images, due to the presence of the strongly correlated speckle noise (see appendix A.3.2), is an even more challenging problem [102]. There exist different approaches for the signal and noise models for ultrasound imaging, depending on the stage of the image formation process (see appendix A.2). These steps are typically divided in: RF signal, envelope signal or log-compressed signal.

The RF signal is hardly available, due to its heavy computational load, although it is the "raw" signal, therefore closer to the physical principles of ultrasound propagation. Within this category, the blind homomorphic deconvolution approach [150, 149] deconvolves the first and second harmonic RF lines (see appendix A.1.3), using a cepstrum method. Since the two signals have poorly correlated speckle patterns, its compounding by averaging tends to remove them. The approach proposed in [12] consists in modeling the ultrasound pulse local polynomial estimation, similar to a wavelet, which permits to robustly estimate the point spread function and deconvolve the image. Michailovich *et al.* [109] estimate the point spread function by wavelet denoising of the RF signal.

The envelope signal noise statistics are considered to be of multiplicative nature [159], following a Rician distribution. Therefore, the use of adapted techniques such as Expectation Maximization [79] is appropriate. Many models have been proposed for the log-compressed signal model, such as additive white Gaussian noise [9], additive colored Gaussian noise [79] or additive with a Fisher Typpett distribution [80]. In the framework of deconvolution, an interesting approach in [132] uses a phantom to characterize the system and compute the inverse filters to perform the deconvolution. However, since the emitted pulse will change in shape as long as it goes through the tissues not necessarily in the same manner as it does through the phantom, the application of this technique is limited. The wavelet decomposition approach has been followed in [74], with results slightly worse than with additive noise. Ultrasound deconvolution for non-destructive testing is examinated in [39], where techniques based on higher order statistics (HOS), subspace methods and neural networks are proposed.

2.2.4 Other specific approaches

Stronger *a priori* constraints can be introduced if there is a solid knowledge about v. The CLEAN algorithm [68] supposes that the original image is composed of point sources, suited for astronomical applications. Local constraints can be introduced by autoregressive (AR) models [108] and non-linear grid filters[156], under a basic hypothesis that the needed information to recover the original image for each pixel is in its vicinity. However, a training stage is needed before their utilization, thus resulting in a dependence on the training set. The NAS-RIF (Non-negativity and Support Constraints Recursive Iterative Filtering) algorithm [92] introduces constraints such as non-negativity and the object size, and can achieve a great performance if the object size is well determined in advance.

In the domain of synthetic aperture radar imaging (SAR), Holmes [70] performed blind deconvolution of images following Poisson statistics using a maximum-likelihood approach. His algorithm consisted in two simultaneous Lucy-Richardson-like deconvolutions to obtain both the original image and the point spread function. Strict positivity is therefore ensured. Holmes also investigated the interest of further explicit constraints such as symmetry or band-limited constraints for the PSF. Lane [94] proposed a variational approach minimizing a penalizing function that measures the deviation from the observed convolution product and the violation of positivity and, possibly, of support constraints for the original image and the PSF, turning into one of the most robust methods.

Another axis of research is to perform a wavelet decomposition of v, in order to better discriminate signal from noise. In general terms, noise is supposed to spread out into all wavelet coefficients, while signal will be concentrated in a few ones. Each decomposition level is deconvolved using different parameters, which can indeed be adaptive, resulting into better results in certain applications [143].

2.3 Multichannel Case

In the single channel case section, we have seen that the frequencies that have been eliminated by the point spread function h can not be recovered. In the multichannel deconvolution framework, it is expected that, from M differently degraded images, it will be possible to recover the original signal v by combining the different parts of the spectrum [82]. By denoting by v_i the ith acquisition through the ith PSF h_i, their relationship is defined as:

$$v_i = h_i * v, \quad i = 1..M \tag{2.19}$$

We will denote the set of all acquisitions v_i by $\{v_i\}$, and the set of filters by $\{h_i\}$. Note that there are M equations and one unknown signal, v. Some authors refer to this scheme as *superresolution* [49], permitting to improve the spatial resolution beyond the size of the PSF by combining different acquisitions.

2.3.1 Exact Deconvolution

One approach to recover v from Equation (2.19) is to find a set of filters $\{d_i\}$ such that:

$$\sum_{i=1}^{M} d_i * v_i = v \tag{2.20}$$

which implies

$$\sum_{i=1}^{M} d_i * h_i = \delta. \tag{2.21}$$

Notice that, differently to the single channel case, the system responses $\{h_i\}$ are inverted through a bank of filters and not a single one, which reduces the ill-posedness of the system. In the z-domain, Equation (2.21) turns into Bezout's identity [23]:

$$\sum_{i=1}^{M} D_i(z) \cdot H_i(z) = 1. \tag{2.22}$$

For one dimensional signals, a necessary condition for Equation (2.22) to have a solution is that $\{H_i\}$ do not have common zeros, that is:

$$\sum_{i=1}^{M} |H_i(z)|^2 \neq 0, \ \forall z. \tag{2.23}$$

As long of one of the filters covers a part of the spectra, exact deconvolution is possible in the noiseless case. Indeed, there exist $\{d_i\}$ of compact support (FIR) if $\{h_i\}$ are filters of compact support (FIR) and if there exist positive constants c_1 and c_2 and a positive integer N such that [23]:

$$\left(\sum_{i=1}^{M} |H_i(z)|^2 \right)^{1/2} \geq c_1 e^{-c_2 |Imz|} \left(1 + |z|\right)^{-N} \tag{2.24}$$

Condition (2.24) is known as strong co-primeness between $\{h_i\}$, and it is a necessary and sufficient condition for the existence of filters $\{d_i\}$ of compact support which are solution of Equation (2.21). For one dimensional signals, coprimeness implies not having zeros in common in the z plane.

For higher dimensions (two and above), the coprimeness of $\{h_i\}$ is a necessary and sufficient condition for the existence of $\{d_i\}$, even if they do have zeros in common in the z plane [65]. In practice, we will consider the condition of not having zeros in common as necessary to obtain an exact deconvolution. This will be, however, a rare case, since frequencies where filter values are low will be dominated by noise.

2.3.2 Least Squares Solution

Similarly to the single channel case, we can consider $\tilde{v}$ that minimizes the squared difference between $\{v_i\}$ and $\{h_i * v\}$, such as:

$$\tilde{v} = \arg\min \sum_{i=1}^{M} \|v_i - h_i * \tilde{v}\|^2 \tag{2.25}$$

In the Fourier domain, the minimum norm solution is given by [23]:

$$\tilde{V}_{LS} = \sum_{i=1}^{M} \frac{H_i^*}{\sum_{j=1}^{M} H_j^* H_j} \cdot V_i \tag{2.26}$$

$\tilde{V}_{LS}$ is only defined if $\{H_i\}$ do not have zeros in common. $\{D_i\}$ defined as $D_i = \frac{H_i^*}{\sum_{j=1}^{M} H_j^* H_j}$ are solution of Equation (2.22).

In order to avoid outliers, other robust norms rather than the L2 can be used on the data fidelity term.

2.3.3 Constrained Least Squares Solution

Adding the constraint (2.10) on $\tilde{v}$, the following energy expression can be derived:

$$E = \sum_{i=1}^{M} \|v_i - h_i * \tilde{v}\|^2 + \lambda \|C \cdot \tilde{v}\|^2 \tag{2.27}$$

For instance, if C is a perfect high-pass filter, $\|C \cdot \tilde{v}\|^2$ would be the energy of $\tilde{v}$ in the band-pass of C. By introducing this term into the energy term, the high frequency content of $\tilde{v}$ is limited. This is typically the case, since deconvolution tends to over-amplify high frequencies.

For a linear C, the minimum norm solution is:

$$\tilde{V}_{RLS} = \sum_{i=1}^{M} \frac{H_i^*}{\left(\sum_{j=1}^{M} H_j^* H_j\right) + \lambda C^* C} \cdot V_i \tag{2.28}$$

Again, the operator C may prevent the zeros in the denominator caused by zeros in the combined spectral response $\sum_{j=1}^{M} H_j^* H_j$, preventing an undesirable noise amplification.

2.3.4 Regularization Techniques

The choice of λ and C in Equation (2.27) represents typically a trade-off between the amount of detail preserved and the amount of noise amplification. They can depend on the *a priori* knowledge about v, on signal and noise statistics, and also on $\{h_i\}$ and on the application specific needs. For instance, if the image has to segmented, a clear separation between the different zones should be preserved.

Regularized image reconstruction has been extensively reviewed in the literature [27, 38, 86, 29, 151, 55] and applied to many domains, such as tomography reconstruction [158].

Linear Operators

One of the most well known example of regularization term is Tikhonov [151]. In the frequency space, it is simply an all-band pass filter, that is, $C(w) = 1$. This guarantees the invertibility of the problem, since it will compensate all the spectral zeros of the system filters. It will effectively block the amplification of high frequencies, but also it will distort the inversion of low frequencies. Other authors [81] have suggested other filters, such as the Laplacian or the inverse Wiener filter of the system.

The regularization term can also be expressed in a general manner in terms of a potential function, ρ, on the differences between neighbors Δ. The regularization energy term will be denoted as:

$$J_{reg} = \sum_{pixels} \rho(\Delta(\tilde{v})) \tag{2.29}$$

In three dimensions, for a 6-voxel neighborhood, the expression (2.29) would be expanded as:

$$\begin{aligned} J_{reg} = & \sum_x \sum_y \sum_z \rho(\tilde{v}(x,y,z) - \tilde{v}(x-1,y,z)) + \\ & \sum_x \sum_y \sum_z \rho(\tilde{v}(x,y,z) - \tilde{v}(x,y-1,z)) + \\ & \sum_x \sum_y \sum_z \rho(\tilde{v}(x,y,z) - \tilde{v}(x,y,z-1)) \end{aligned} \tag{2.30}$$

The expression for other neighborhood configurations is straight-forward.

The Tikhonov regularization term is:

$$J_{Tikhonov} = \|\Delta(\tilde{v})\|^2 \tag{2.31}$$

With this notation, we are explicitly describing the local relationship between the pixels in the image. Neighbor pixels which are very different ($\Delta(\tilde{v})$ is large) will have a large contribution to the energy term. On the contrary, neighbors which have the

same intensity level ($\Delta(\tilde{v}) = 0$) will not increment the energy term. Minimizing the Tikhonov regularization term will induce to smooth the image, make the intensity values similar among neighbors.

Edge-preserving regularization

Typically, edges give valuable information that should be retained in the reconstructed image. Regularization terms such as Tikhonov do not preserve those edges, since they smooth the image regardless of the position of the pixels. For these reasons, edge-preserving regularizing techniques have been developed [27, 38]. These techniques include a discontinuity in the potential functions to avoid the smoothing of edges. For instance, the Huber [25] function introduces a threshold value (α), over which the difference between the neighbors are treated with a linear potential rather than a quadratic one. It is defined as:

$$\rho(x) = \begin{cases} x^2, & \text{for } |x| \leq \alpha \\ 2\alpha\,|x| - \alpha^2, & \text{for } |x| > \alpha \end{cases} \tag{2.32}$$

and its derivative is:

$$\rho'(x) = \begin{cases} 2x, & \text{for } |x| \leq \alpha \\ 2\alpha\text{sign(x)}, & \text{for } |x| > \alpha \end{cases} \tag{2.33}$$

Figure 2.2 shows some examples of potential functions ρ. Other regularization techniques include total variation (TV) [130], and the more recently introduced bilateral filter [152]. The median filter can also be used as a regularization operator. Loupas *et al.* [102] proposed an adaptive median filter adapted to speckle for ultrasound imaging.

Anisotropic diffusion [122] can also be used as regularization term. Indeed, Black *et al.* [26] showed the relationship between robust estimators and anisotropic diffusion, which turns out to be equivalent for some particular potential functions. Particularly, the diffusion equation proposed in [122], as:

$$\frac{\partial v(\mathbf{x}, t)}{\partial t} = \nabla \cdot (g(\|\nabla v\|)\nabla v) \tag{2.34}$$

where t represents the artificial time scale, $\nabla \cdot (\bullet)$ the spatial divergence operator, $\nabla(\bullet)$ the spatial gradient, $\|\nabla(\bullet)\|$ the gradient magnitude, and $g(\bullet)$ the edge-stopping function. Using one of the functions originally proposed in [122], defined as:

$$g(x) = \frac{1}{1 + \frac{x^2}{2\alpha^2}} \tag{2.35}$$

is equivalent to minimize the regularization term defined in Equation (2.29) using the Lorentzian function [26], defined as:

$$\rho(x) = \log\left[1 + \frac{1}{2}\left(\frac{x}{\alpha}\right)^2\right] \tag{2.36}$$

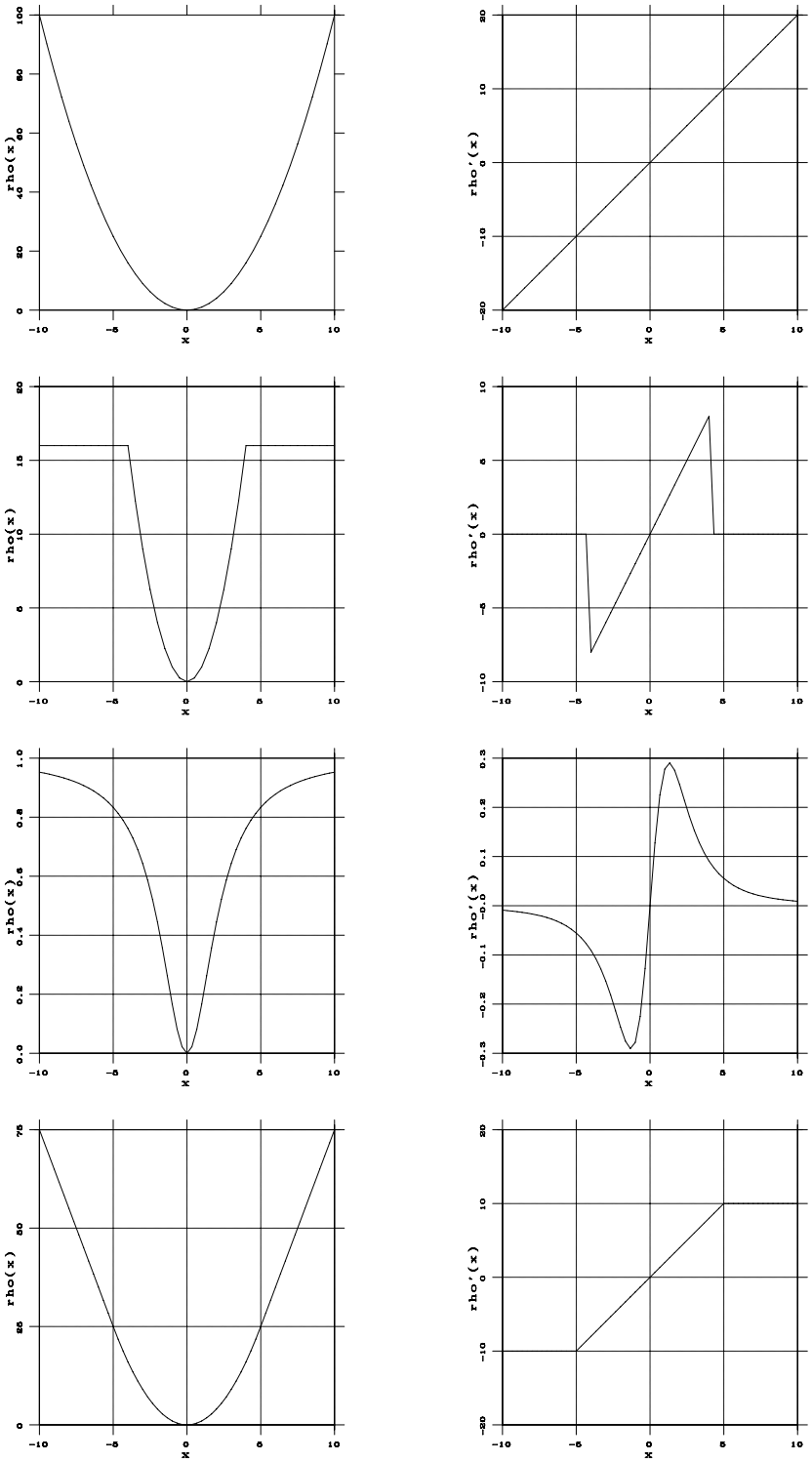


Figure 2.2: Some examples of robust estimators. Left: ρ function. Right: its derivative. From top to bottom: quadratic (Tikhonov) [151], truncated quadratic [115], Geman-McLure [55], Huber [25].

and its derivative

$$\rho'(x) = \frac{2x}{2\alpha^2 + x^2} \tag{2.37}$$

Elad [46] showed the relationship between bilateral filtering and anisotropic diffusion. Farsiu *et al.* [49] used the bilateral filter as regularization term in the superresolution context. Several adaptations of anisotropic diffusion for spatial filtering have been recently proposed for ultrasound imaging [89, 148, 110], and for synthetic aperture radar (SAR) [166], which exhibits the same problems than speckle for ultrasound imaging.

2.3.5 Iterative Optimization Algorithms

The energy in Equation (2.27) can also be minimized by iterative techniques, which avoid to compute explicitly the inverse filters and allows the use of non linear C operators. In this section we develop their expression for the multichannel case.

Steepest Descent

The steepest descent technique consists in modifying the estimate $\tilde{v}$ in the direction of stronger variation, that is the opposite direction of the energy gradient. Denoting by $\tilde{v}^k$ the solution $\tilde{v}$ at the kth iteration, and r^k the residual that updates $\tilde{v}$ with a step β, we have:

$$\tilde{v}^0 = \frac{1}{M}\sum_{i=1}^{M} \tilde{v}_i$$

$$r^k = -\frac{1}{2}\nabla E(\tilde{v}^k)$$

$$\tilde{v}^{k+1} = \tilde{v}^k + \beta r^k$$

This technique converges toward the nearest local minimum, therefore a proper initialization is critical. If the energy term is convex, the algorithm converges toward the global minimum for any initialization.

For the multichannel energy defined in Equation (2.25) (without regularization), we have, in the spectral domain:

$$R^k = \sum_{i=1}^{M} H_i^* \left(V_i - H_i \tilde{V}\right) \tag{2.39}$$

Note that the residual is a sum of terms for each acquisition, each term contributing only to the pass-band of H_i. Therefore, all out-of-the band noise is rejected and common band-pass frequencies are averaged. At the kth iteration, the solution $\tilde{V}^k$ is:

$$\tilde{V}^k = \beta \left(\sum_{i=1}^{M} H_i^* V_i\right) \sum_{l=0}^{k} \left(1 - \beta \sum_{j=1}^{M} H_j^* H_j\right)^k \tag{2.40}$$

$\tilde{V}^k$ converges for $k \to \infty$ to $\tilde{V}_{LS}$ in Equation (2.26) if the right-hand side series of Equation (2.40) converges. This occurs if $\left(1 - \beta \sum_{j=1}^{M} H_j^* H_j\right) < 1$, which happens if $\sum_{j=1}^{M} H_j^* H_j(w) \neq 0,\ \forall w$ ($\{h_i\}$ have no common zeros) and for:

$$0 < \beta < \frac{2}{\max \sum_{j=1}^{M} H_j^* H_j}$$

Note that β should be M times smaller than in the single channel case.

Notice that the right hand of Equation (2.40) does not depend on the input signals V. Therefore, for known H, this part of the filter can be computed off-line.

For the energy defined in Equation (2.27) (with regularization), we have, in the spectral domain:

$$R^k = \sum_{i=1}^{M} H_i^* \left(V_i - H_i \tilde{V}\right) - \lambda C^* C \tag{2.41}$$

and

$$\tilde{V}^k = \beta \left(\sum_{i=1}^{M} H_i^* V_i\right) \sum_{l=0}^{k} \left(1 - \beta \left(\sum_{j=1}^{M} H_j^* H_j + \lambda C^* C\right)\right)^k \tag{2.42}$$

Equation (2.42) converges if $\sum_{j=1}^{M} H_j^* H_j + \lambda C^* C \neq 0,\ \forall w$. In this case, the operator C may prevent the common zeros in $\{H_i\}$. β should be within the range:

$$0 < \beta < \frac{2}{\max \sum_{j=1}^{M} H_j^* H_j + \lambda C^* C}$$

Conjugate Gradient Descent

The conjugate gradient method uses conjugate directions instead of the local gradient to update the image estimation. If the energy function has the shape of a long, narrow valley, the minimum is reached in fewer steps than would be the case using the method of steepest descent, which exhibits a linear rate of convergence.

The solution $\tilde{v}$ at the kth iteration is modified in the direction p^k such as [18]:

$$\tilde{v}^0 = \frac{1}{M} \sum_{i=1}^{M} \tilde{v}_i$$

$$p^0 = r^0$$

$$r^k = -\frac{1}{2} \nabla E(\tilde{v}^k)$$

$$p^k = r^k + \frac{\left\langle r^k, r^k \right\rangle}{\left\langle r^{k-1}, r^{k-1} \right\rangle} \cdot p^{k-1}$$

$$\tilde{v}^{k+1} = \tilde{v}^k + \alpha p^k$$

Again, a proper initialization is required to avoid that the algorithm is trapped in a local minimum.

2.4 Multiview Case

2.4.1 Multiview Acquisitions

The multiview case is a particularization of the multichannel case, where the different point spread functions are related geometrically. We denote by h_i the point spread function of the ith acquisition. They are related by a certain transformation T, such as $h_i(\mathbf{x}) = h_j \circ T(\mathbf{x})$. In fact, we will denote the generic point spread function by h, and the transformation associated at each acquisition by T_i. The ith point spread function therefore is:

$$h_i(\mathbf{x}) = h \circ T_i(\mathbf{x}) \tag{2.44}$$

In practice, the geometrical transformations will typically be rotations and translations. In three dimensions, they can be conveniently expressed in homogeneous coordinates as:

$$T_i = \begin{pmatrix} r_i^{00} & r_i^{01} & r_i^{02} & d_i^x \\ r_i^{10} & r_i^{11} & r_i^{12} & d_i^y \\ r_i^{20} & r_i^{21} & r_i^{22} & d_i^z \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.45}$$

where $R_i = \{r_{jk}\}$ represents the centered rotation matrix, and $D_i = [d_x, d_y, d_z]^T$ represents the translation vector.

For purely translational transformations ($R_i = \mathbf{I}$), the Fourier transform of the point spread function h_i has the same spectral power that h, but with a phase shift, such as:

$$h_i(\mathbf{x}) = h(\mathbf{x} + D_i) \Rightarrow H_i(\mathbf{w}) = H(\mathbf{w}) e^{-jD_i\mathbf{w}} \tag{2.46}$$

When a rotation is involved, the Fourier transform is rotated with the same angle:

$$h_i(\mathbf{x}) = h(R_i(\mathbf{x})) \Rightarrow H_i(\mathbf{w}) = H(R_i(\mathbf{w})) \tag{2.47}$$

This is one of the different ways to perform spatial compounding in ultrasound echography, where the tissue is insonified with different angles [162]. In this case, while rotating the ultrasonic probe, thus the point spread function, different parts of the spectra will be insonified. In this way, frequencies not accessible from one acquisition will be gathered by another one. Figure 2.3 shows 2D examples of synthetic point spread functions along with their Fourier transforms. When rotated, the Fourier transform follows the same rotation as the space domain point spread function.

2.4.2 Spectral Coverage

The different acquisitions cover differently the spectral domain. Acquisitions will be *redundant* at the frequencies which are in the bandpass of two or more views. This redundancy will be used to reduce the noise figure of the system, exploiting the complementary information of each view, since ideally noise is uncorrelated between the

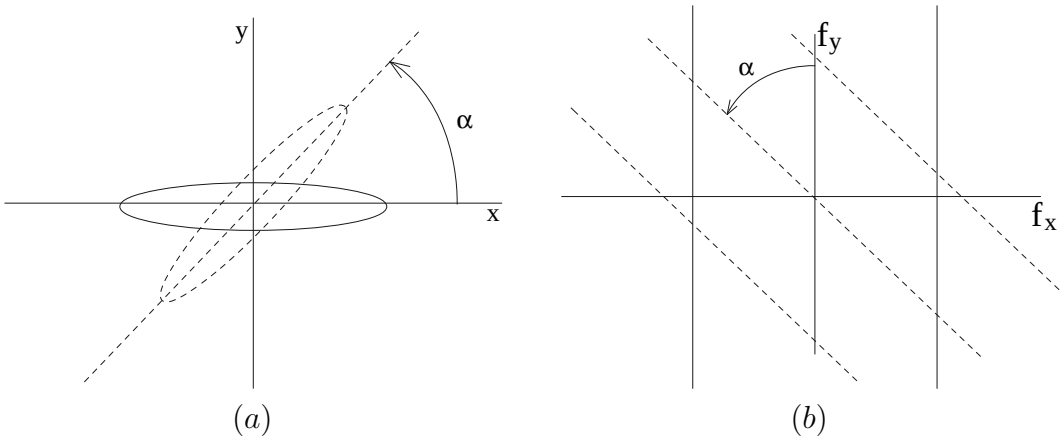
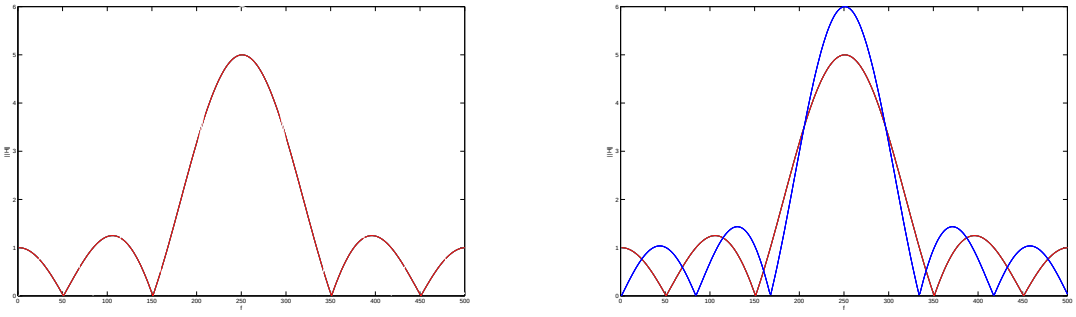


Figure 2.3: Schema of two rotated point spread functions: a) in the spatial domain and b) their Fourier transform. A rotation in the space domain implies a rotation in the frequency domain of the same angle.

views while signal is correlated. On the other hand, artifacts like shadowing (see Appendix A.3.5), may induce conflictual information at these parts of the image.

Frequencies which are covered by just one of the views are subject to the same signal to noise problems as in single channel deconvolution, thus the need of regularized methods, at least at these frequencies.

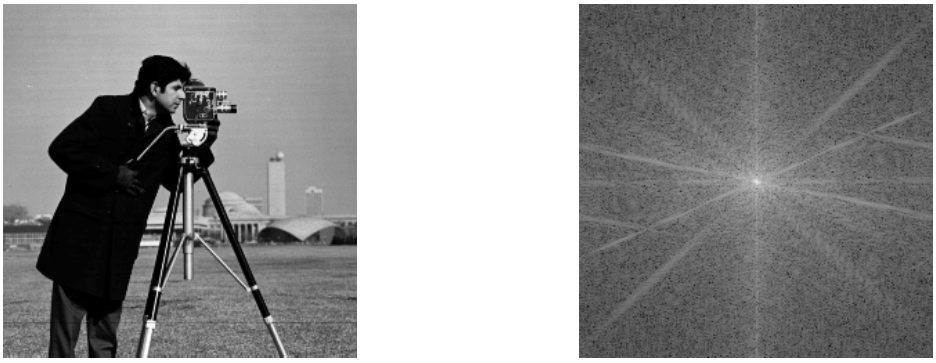
As seen in Section 2.3.1, exact reconstruction of the original image will only be possible if the different H_i do not have zeros in common, that is, if they cover the whole spectral space. In other words, if there are no incomplete frequencies. Figure 2.4 illustrates this principle. In the single channel case shown at left, there are several frequencies where the response of the point spread function is zero. On the other and, the figure on the right shos how another different acquisition might give complementary information at these frequencies. In this case, in the absence of noise, exact deconvolution is possible.



a) single-channel case b) two-channel case

Figure 2.4: a) Spectral response of a low-pass PSF, b) Spectral response of a larger low-pass PSF, along with a). The frequencies canceled by one of the channels can be recovered on the other channel.

However, not always the whole spectral plane is covered. Figure 2.5 shows the original example image along with its Fourier transform. Figure 2.6 shows four simulated views of the original image with four point spread functions rotated by 0, 45, 90 and 135 degrees respectively. It can be seen how each of the acquisitions preserves some of the features of the original image, which in turn can be appreciated in the part of the spectra that it covers. For example, in the view blurred vertically (Figure 2.6.b) the middle stick of the tripod of the cameraman is clearly visible. In the view blurred horizontally (Figure 2.6.h), this part is hardly distinguishable. On the other hand, in this view the roof of the tall buildings of the background can be discriminated, while this is not possible in the vertically blurred view. The same principle applies to the other views, such as the diagonal stick of the tripod only present at the acquisition on 135 degrees (Figure 2.6.k).



(a) (b)

Figure 2.5: a) Original image "cameraman", and b) its Fourier transform, log-compressed for displaying purposes.

Figure 2.7 shows a real example of 2D ultrasonic images acquired at different angles. As in Figure 2.6, each acquisition retains certain features at certain angles in function of the PSF. Also, it can be appreciated how the speckle patterning differs between them, specially when angle differences are large. Other artifacts, such as clutter (see Appendix A), show different patterning between the different acquisitions. By combining these different images, a closer representation of the original tissue is obtained. Current commercial systems, such as SonoCT^{tm} (Philips, Best, The Netherlands), perform an averaging of these views to combine them, also shown in Figure 2.7.

The goal of multiview deconvolution is to restore the original image from these acquisitions. We will evaluate the techniques presented in this chapter, reconstruction from anisotropically degraded images, in Chapter 3. where we will introduce the techniques to estimate the point spread function. The estimation of the point spread function is a required step to deconvolve these images.

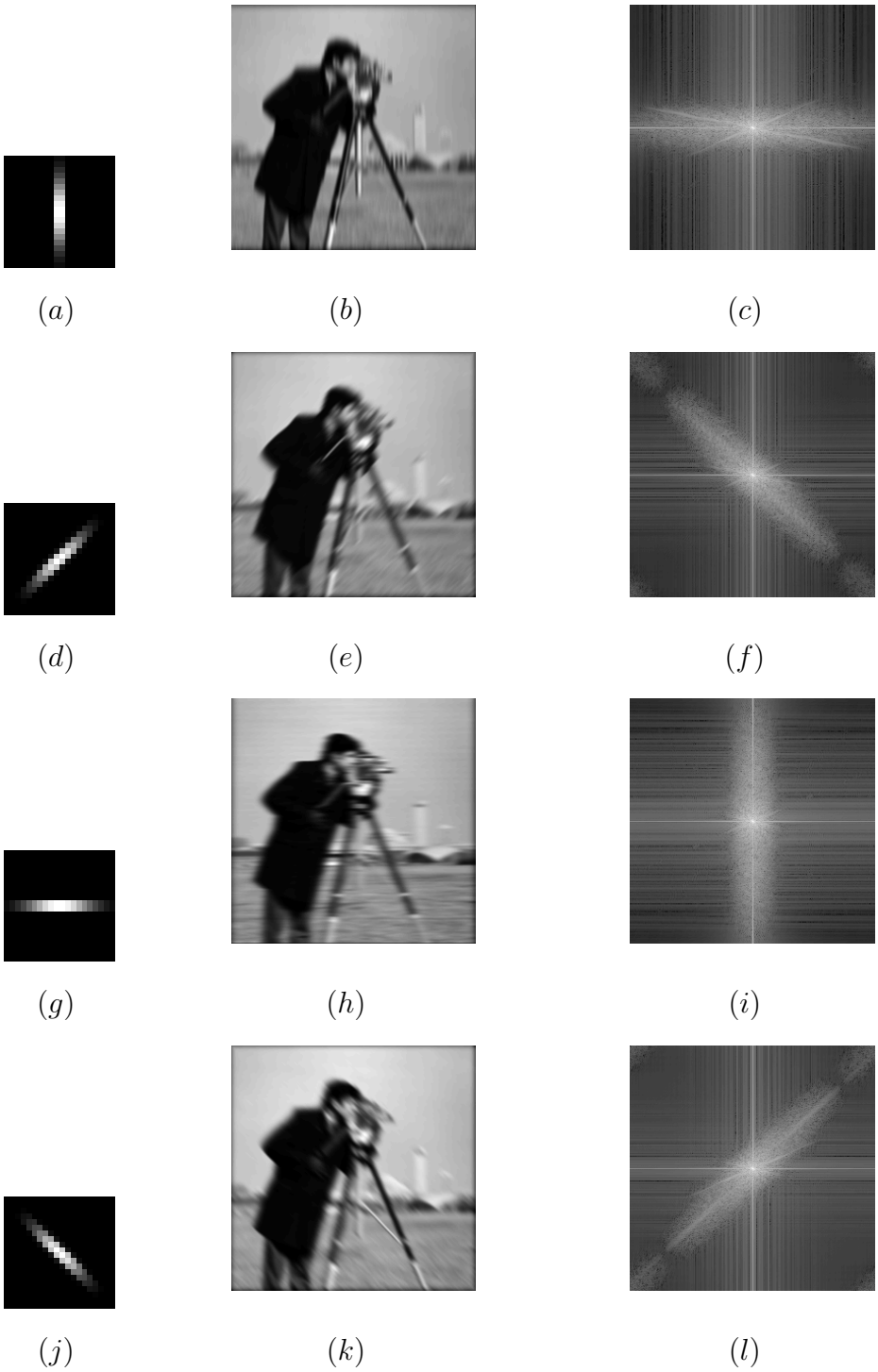


Figure 2.6: Synthetically blurred images. Left, point spread functions. Middle, blurred images. Right, Fourier transform of blurred images. From top to bottom, simulated acquisitions at 0, 45, 90 and 135 degrees respectively.

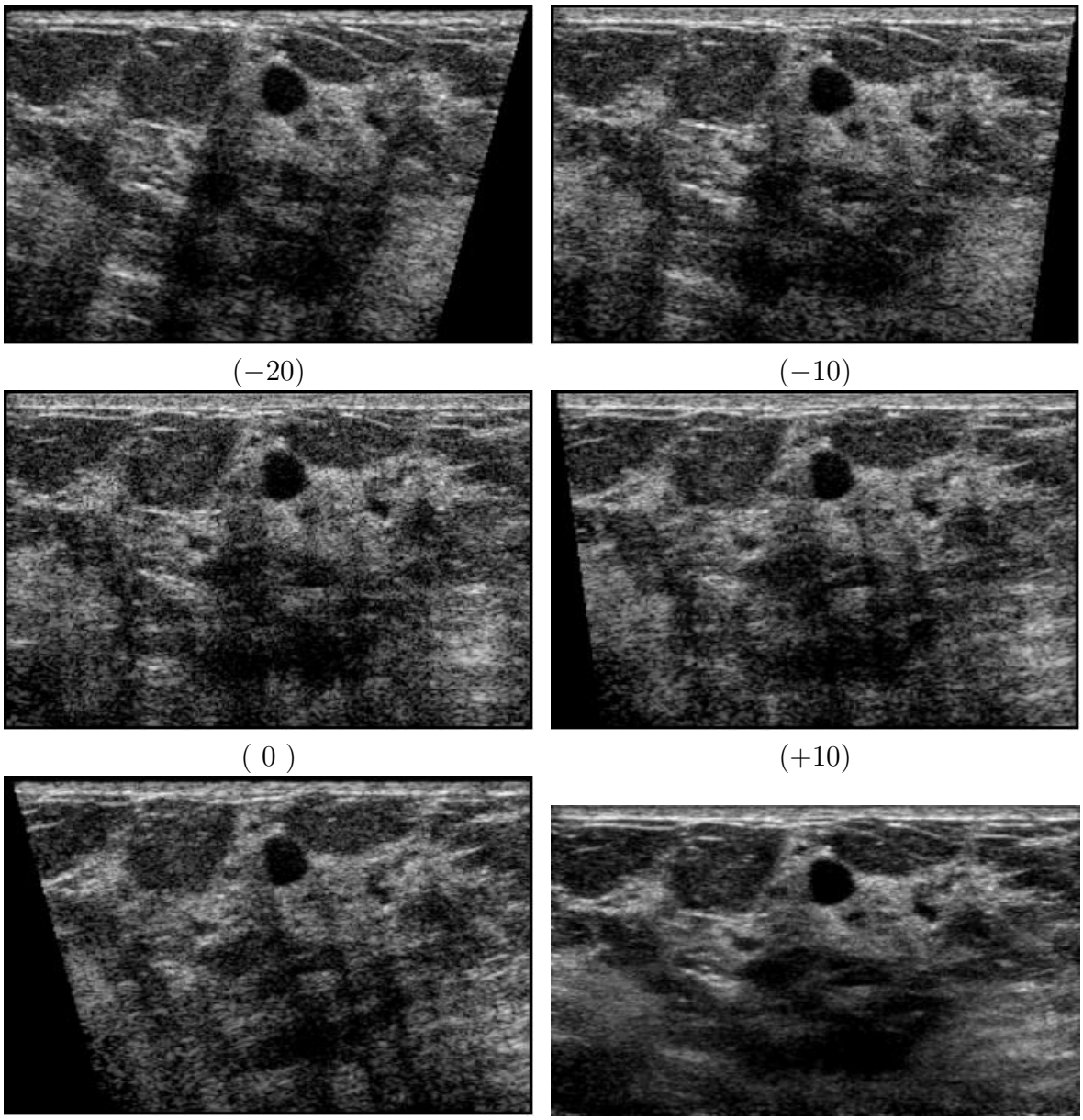


Figure 2.7: *In vivo* ultrasonic acquisitions at different insonification angles. The different acquisitions show different speckle, clutter and shadow patterns. Their combination results in an improved image quality result.

2.5 Conclusions

We reviewed the most common techniques for single channel deconvolution. In general, deconvolution is an ill-posed problem due to spectral zeros and the presence of noise. To overcome this difficulty, *a priori* information about the searched solution is provided in form of regularization, which constrains the solution space. It has the form of a linear operator constraint for the deterministic algorithms and the form of probability density function in the Bayesian framework. However, exact deconvolution is limited to very few applications for which the solution space is small and can be well characterized. In general, at the frequencies were data have been filtered the original will not be recovered.

In the multichannel case, the ill-posedness of the problem is reduced. In fact, as long as one of the acquisitions contains a part of the spectrum, it can be recovered as in the single-channel case. However, if a part of the spectrum is not covered by any acquisition, the problem is the same as for the single-channel case. In this sense, multichannel deconvolution deals about how to combine the different views rather to how to deconvolve them, which is done by the same principles as in single-channel deconvolution.

In particular, for the case where the different acquisitions are related geometrically, that we denote as multiview case, each view covers a part of the spectrum directly related to its geometrical transformation. In this way, acquisitions with the same instrument but from different points of view can complement each other toward a better image restoration. We denote as spectral coverage the part of the spectrum insonified by the set of all acquisitions. Each acquisition captures features in a certain orientation, and the proposed deconvolution technique captures all these features into a single combined volume.

In all cases, regularization is needed in order to avoid noise amplification. Indeed, it is a convenient step to introduce *a priori* knowledge on the signal in order to reduce the space of solutions of the proposed inverse problem is promising. Classical regularization techniques tend to blur edges, thus the introduction of edge-preserving regularization operators via robust estimation. The estimation of the hyperparameters is still an open question. Further improvements could be obtained by scale-space decomposition, e.g. Laplacian Pyramid [34] or wavelet decomposition [105], and adapting the parameters of regularization in function of the scale.

In this chapter, the point spread function of the system was supposed to be known. Indeed, no specific particularization on ultrasound imaging was done. In the next chapter we tackle the problem of estimating the point spread function, and demonstrate on synthetic data the performance of the deconvolution techniques developed so far. This allows us to choose a particular regularization approach. In Chapter 4 we give more details on ultrasound specificities and the used noise model. Application to real ultrasound data is left for Chapters 5 and 6.

CHAPTER 3

Blind Multiview Deconvolution

3.1 Introduction

In Chapter 2 we introduced the theoretical basis to deconvolve an image from differently degraded images. The point spread functions of the system were supposed to be known. Unfortunately, this is not the case in ultrasound imaging, moreover, the estimation of the PSF only from the data itself is not a trivial problem. Typically, techniques that estimate the PSF in addition to the original image are called *blind* deconvolution techniques. In this chapter we present a new technique to estimate the point spread functions from the different acquisitions by using *a priori* information of our system. This includes the geometrical relationship between the different PSF but also shape constraints of these point spread functions.

As a reminder, the blind deconvolution problem consists in obtaining v from v_d without the knowledge of h, in the convolution equation:

$$v_d = h * v \tag{3.1}$$

We start in Section 3.2 with a review of the state-of-the art blind deconvolution techniques, both for single channel and for multichannel cases. The approaches are radically different, since the multichannel strategy makes use of the diversity of the different channels, which is not available in the single channel case. From the multichannel techniques, we chose the subspace technique as starting point in Section 3.3. The main problems of this technique are the robustness to noise and the dependency on support estimation. To improve these points, we introduce the geometrical constraints of the multiview case, where the different PSFs are rotated versions of a common kernel. In this way, the number of unknown variables is reduced. In order to further increase its performance, up to realistic figures of noise, shape constraints are also introduced, thus reducing the estimation of the PSF to a parameter optimization problem. The results of these techniques along with those of Chapter 2 are evaluated on synthetic data to restore images from anisotropically degraded views in Section 3.4. Finally, conclusions are discussed in Section 3.5.

3.2 Blind Deconvolution

3.2.1 Single Channel Case

Several techniques have been developed to estimate the point spread function from an image itself, although it is not always possible to obtain satisfactory results. Blind deconvolution is, in general, an undetermined problem, having one equation and two unknown images (h and v) in the noiseless case, and three in the noisy case (also n). This leads to the need of additional constraints to recover the original image.

If some properties of the image, the point spread function, or the noise are known *a priori*, these can be used to reduce the ill-posedness of the problem. Typical examples are non-negativity or finite support. Statistical properties of the image and noise are also used in some techniques. If the physical nature of the degradation is known, parametric models of the PSF can be used. Some authors designate as *myopic* or *partially blind* the techniques that use some knowledge about the PSF. We will not use this terminology here, and we will use the term *blind* in all cases.

However, there are some intrinsic ambiguities in blind deconvolution systems, due to the convolution properties. First of all, the image and the PSF must be irreducible, which means that any of these cannot be expressed as the convolution of two or more components. For instance, if $h = h_1 * h_2$, $v_d = h * v = h_1 * h_2 * v$, an ambiguity occurs in deciding which components belong to the image and which to the PSF. A Gaussian PSF, for instance, is reducible, since it can be expressed as a convolution of two Gaussians. In this case, the support estimation will be critical to capture the correct PSF.

There is also an intrinsic shift ambiguity between the image and the PSF, since $v_d(\mathbf{x}) = h(\mathbf{x} + \Delta\mathbf{x}) * v(\mathbf{x}) = h(\mathbf{x}) * v(\mathbf{x} - \Delta\mathbf{x})$. In other words, an arbitrary shift in the PSF and the opposite shift in the image are undistinguishable. Therefore, a wrong estimation of the PSF may induce a shift in the reconstructed image. Similarly, a scaling factor is also ambiguous, since $v_d = h * v = (1/\alpha)h * \alpha v$. Shifting and scaling ambiguities are not typically major issues, since they might be solved with information about the localization of the image or the PSF, and including other *a priori* information, such as mean value preservation ($\sum_{-\infty}^{\infty} h = 1$) or by normalizing the final reconstructed image. However, the support estimation of the PSF appears typically as a critical issue. If it is underestimated, no solution is found, and if it is overestimated, spurious effects appear.

Algorithms can be divided into two main groups, those that estimate the PSF before performing deconvolution, and those that estimate the PSF and the original image simultaneously. We will briefly discuss some of the algorithms here, pointing the reader to [66, 91] for a complete review on the subject.

Stand-alone PSF Estimation

Chang *et al.* [37] proposed the frequency domain zeros method. Given the convolution equation in the frequential domain, $V_d = H \cdot V$, it is easy to see that zeros in this domain are due to H, V or both. The main idea of this technique is to identify which zeros are

due to H, and recover H from those. To identify the zeros, the image is partitioned in small windows, large enough to contain the support of h. Secondly, to recover H from its zeros, typically a parametric form is needed. The main drawback of this method is that it is not suited for degradations which do not have zeros, e.g. Gaussian functions, which is typically the case in ultrasound echography.

If features such as points or edges can be identified in the blurred image, with assumptions on how these features would look like in the true image, an attempt to recover the PSF can be made [19]. Indeed, if the PSF has a known parametric form, these features can be used to find out the parameters. Typically these techniques are used in applications such as astronomy, where single dots with uniform background can be easily identified. Otherwise, in other applications such as ultrasound echography, it is difficult to guarantee the existence of sharp edges.

Simultaneous image/PSF estimation

There is another class of algorithms which evaluate the PSF and the original image at the same time. The iterative blind deconvolution (IBD) [17] is one of the most popular ones within this class. It is an iterative algorithm which alternates the estimation of the PSF and the image, imposing constraints both in the spatial and frequential domains. The initial guess for the reconstructed image is transformed to the frequency space, where constraints are applied (e.g. band-limiting filters). In the frequency space, the PSF is estimated from the current reconstructed image and the degraded image, through a Wiener-like filter. The PSF is transformed into the space domain, where constraints are applied (e.g. non-negativity, replacing negative values with zeros). The PSF is transformed to the frequency space, constraints are applied, and the reconstructed image is estimated. The iteration goes on until convergence or a user-specified maximum number of iterations is reached. The algorithm has a low computational complexity but its convergence is quite uncertain and shows a significant dependence on parameters, turning into a lack of reliability. The "Non-negativity And Support constraints Recursive Inverse Filter" (NAS-RIF) [92] algorithm overcomes the convergence problems of the IBD, but at the cost of dealing with objects of known size over uniform backgrounds.

Other approaches minimize directly an energy term (e.g. $\|h * v - v_d\|^2$), both with respect to h and v, using different methods, such as Generalized Cross-Validation (GCV) [127] or Simulated Annealing (SA) [107]. Besides the high computational cost of such methods, the ill-posedness with respect to both h and v limits the reliability of such methods.

To reduce the number of unknowns, Auto Regressive - Moving Average (ARMA) models have been proposed. The image is modeled as an autoregressive (AR) process, and the PSF as a moving average (MA) process. Many techniques have been proposed to estimate the ARMA parameters, such as Maximum Likelihood [93]. However, these models are adequate for images which are smooth and homogeneous, failing when the original image contains edges.

The zero sheet method [95] relies on the fact that if the PSF and the image are irreducible, the z-transform of the degraded image $V_d(\mathbf{z}) = H(\mathbf{z}) \cdot V(\mathbf{z})$ can be factorized

into the PSF and the original image. Despite the conceptual interest of this method, the computational complexity and specially the lack of robustness against noise make this method impractical for real systems.

3.2.2 Multichannel Blind Case

The problem of blind deconvolution in the multichannel case differs completely from the single channel case, because the different filters may be estimated from their complementary information. In general, it is still however an undetermined problem, having M equations and $M+1$ unknown images (h_i and v) in the noiseless case, and $2M+1$ unknown images (plus n_i) in the noisy case. Assumptions such as modeling the PSF as a FIR filter of support much smaller than the image reduces dramatically the dimensionality of the problem.

There exist mainly three approaches for blind multichannel deconvolution:

1. estimate the blurs, then deconvolve the image (e.g. [65])

2. compute directly deconvolver filters (e.g. [56])

3. compute directly the restored image (e.g. [163, 119])

The first approach offers more control over the restoration process, due to its ill-posed nature. It is less obvious in the second and third approaches to introduce regularizing methods or the like. Therefore, we will focus on the first approach.

Several techniques have been proposed for PSF estimation, specially in the domain of 1D channel identification. A review on 1D multichannel blind estimation techniques techniques can be found in [153]. Although the borders between these techniques are not very clear, they may be classified into Maximum Likeklihood (ML) and Moment methods, both in deterministic or stochastic approaches, depending on the chosen signal model. The subspace method is a subcategory of the moment methods, and will be treated in detail within this section.

Once the filters are estimated, many techniques are possible to reconstruct the original image, as shown in Chapter 2. If FIR deconvolvers are computed [65, 56], deconvolution can be achieved by a single FIR filtering, which is less computationally complex than the least-squares reconstruction, in detriment of noise robustness. Another important branch of approaches is to alternate the estimation of the PSF and the reconstruction. If the estimation of the PSF depends on the reconstructed image, this could improve the estimation as the reconstructed image converges toward the original image. Blind multichannel deconvolution is approached in [141] as an energy minimization of three terms, namely: data fidelity, regularization and PSF estimation using subspace methods. The Bussgang algorithms (e.g. [21]) falls also within this class of alternate algorithms. Panci *et al.* [121] apply different constraint filters depending on the structure of the image (edges, points, ridges, etc.).

Channel Disparity

In order to estimate the blurs from the different channels, they must exhibit certain disparity. This condition is expressed [120] as *factor* (or *weak*) co-primeness between them. A set of FIR filters $H_i(\mathbf{z})$ are factor co-prime if they do not have factors in common, that is, if and only if the greatest common divisor is a scalar, i.e., $H_i(\mathbf{z}) = C(\mathbf{z})H_i'(\mathbf{z}), \forall i = 1 \dots M$ is true only for $C(\mathbf{z})$ a scalar factor. Weak co-primeness is equivalent to the condition that filters are irreducible.

While in 1D *factor* co-primeness implies *zero* co-primeness (introduced in section 2.3.1, as not having a common zero), in 2D, factor co-primeness is a much less restrictive condition. This implies that blur identification is possible even if the filters have common zeros. Indeed, theoretically, two 2D FIR filters are generically factor co-prime [65], which means that blur identification is possible if only two channels are available. Moreover, zero co-primeness is also generically satisfied if there are three 2D FIR filters [65], which leads to consider that exact deconvolution in the absence of noise is generically possible if three channels are available. However, due to the presence of noise and the fact that real systems may not be modeled exactly as FIR filters, these hypotheses may be easily violated in real systems.

Subspace techniques

Considered by some authors as part of moment methods, the subspace methods have been widely used for blind multichannel deconvolution. One of the first to introduce it for 1D signals were Gürrelli *et al.* [62], proposing eigenvector-based algorithm (EVAM) to estimate the blur functions. Many other works have been developed within the subspace approach for 1D signals (e.g. [154, 112, 10]). Giannakis *et al.* [56] extended this technique to 2D, and similar works have been presented for instance in [65, 119, 141].

Subspace methods provide a closed form for the blurs, while the ML methods have to be solved iteratively. Indeed, in the deterministic version, subspace methods make no assumptions about the original image. However, they are not very robust to noise and need an accurate estimate of the PSF support, which is often difficult. We will see in the next section how the multiview constraint improves the performance of the algorithm with respect to these points.

Subspace methods rely on the cross-channel relationship, which states that:

$$h_i * v_j - h_j * v_i = 0 \quad \forall i, j \; i \neq j \tag{3.2}$$

which is easily deduced from the commutativity of the convolution operator, since $h_i * v_j = h_i * h_j * v = h_j * h_i * v = h_j * v_i$. For example, v_1 represents the original image v blurred with the filter h_1, and if v_1 is further blurred with h_2, it will be equal to v_2 blurred with the filter h_1.

If the co-primeness condition of filters h_i are met, solutions of:

$$v_i * g_j - v_j * g_i = 0 \quad \forall i, j \; i \neq j \tag{3.3}$$

have the form

$$g_i = \begin{cases} h_i * k & \text{if } m_g \geq m_h \vee n_g \geq n_h, \\ \alpha h_i & \text{if } m_g = m_h \vee n_g = n_h, \\ 0 & \text{if } m_g < m_h \wedge n_g < n_h. \end{cases} \tag{3.4}$$

where (m_h, n_h) is the size of the support of filters h_i, (m_g, n_g) is the size of the support of filters g_i, k is some factor of size $(m_g - m_h + 1, n_g - n_h + 1)$ and α is a scalar. Notice that Equation (3.3) does not depend on the original image but only on the acquired images and the filters. If the filters were not co-prime, the common factor would be canceled out in Equation (3.3).

If noise is present, the right side of Equation (3.3) is no longer zero but some measurement of noise. Then the set of Equations (3.3) can be solved as a least-squares minimization for g_i, that is:

$$\{g_i\} = \arg\min \sum_{1 \leq i < j \leq M} \|v_i * g_j - v_j * g_i\|^2 \tag{3.5}$$

Support Estimation

As expressed in Equation (3.4), support estimation is critical for successful blur identification, since underestimation has no solution, and overestimation exhibits spurious effects. Harikumar *et al.* [65] propose the residual-based technique, based on the idea that the residual of Equation (3.5) will be minimal for the correct support size. An iterative approach reconstructs the blurs for different support sizes to minimize the residual. Giannakis *et al.* [56] propose a technique to estimate the blur support from the rank of $\{g_i\}$. Given that the solution of Equation (3.3) is unique for the correct support size and has no solution for underestimations of it, the rank of the solution matrix is studied through eigen-vector decomposition. In the presence of noise, a threshold is set to consider the effectively nonzero singular values to estimate the rank. In general these algorithms depend on the noise model used and the parameter estimations made.

Edge effects

The discussion followed up to now assumes that the complete result of the convolution is available. This is seldom the case, since at the borders only partial information is available. We will not discuss the issues raised by this fact (see [65] for a treatment of this problem), and we limit ourselves to the reconstruction of images within a margin from the edges where full data are available.

3.3 Subspace Techniques

Our interest on subspace techniques is three-fold:

1. the general principle does not depend on a particular image or noise model,

2. the solution has a closed-form for the filters,

3. it is easy to introduce the multiview constraint.

We will detail the algorithm in two dimensions, without loss of generality. Let prime denote transpose, lower bold case vectors and upper bold case matrices. In order to express the relationship between h_i and v_i more conveniently, we will use the matrix form, similar to [56]. Let (L_1, L_2) be the minimum support size of filters h_i, and (N_1, N_2) the size of images v_i. We express the PSF in a column vector of size $(L_1+1)(L_2+1) \times 1$ as:

$$\mathbf{h}_i \equiv [h_i(0,0) \dots h_i(0,L_2); \dots; h_i(0,0) \dots h_i(L_1,L_2)]' \tag{3.6}$$

which is the lexicographically ordered two dimensional PSF. Similarly, one image column of size $(L2+1)$ is expressed as a row vector of size $1 \times (L2+1)$ as:

$$\mathbf{v}_j(n_1;n_2) \equiv [v_j(n_1,n_2) \dots v_j(n_1,n_2-L_2)] \tag{3.7}$$

Then, the two dimensional convolution of the image $\mathbf{v}_j$ with the PSF $\mathbf{h}_i$ can be expressed as:

$$\begin{bmatrix} \mathbf{v}_j(N_1-1;N_2-1) & \cdots & \mathbf{v}_j(N_1-1-L_1;N_2-1) \\ \vdots & \vdots & \vdots \\ \mathbf{v}_j(N_1-1;L_2) & \cdots & \mathbf{v}_j(N_1-1-L_1;L_2) \\ \hline \vdots & \vdots & \vdots \\ \hline \mathbf{v}_j(L_1;N_2-1) & \cdots & \mathbf{v}_j(0;N_2-1) \\ \vdots & \vdots & \vdots \\ \mathbf{v}_j(L_1;L_2) & \cdots & \mathbf{v}_j(0;L_2) \end{bmatrix} \times \begin{bmatrix} h_i(0,0) \\ \vdots \\ h_i(0,L_2) \\ \hline \vdots \\ \hline h_i(L_1,0) \\ \vdots \\ h_i(L_1,L_2) \end{bmatrix} = \mathbf{V}_j\mathbf{h}_i \tag{3.8}$$

where the matrix $\mathbf{V}_j$ has dimensions $(N_1-L_1)(N_2-L_2) \times (L_1+1)(L_2+1)$, and the convolution result $\mathbf{V}_j\mathbf{h}_i$ is a column vector of size $(N_1-L_1)(N_2-L_2) \times 1$, where each row represents a pixel of the convolved image. The cross relation approach in Equation (3.3) can be stated in this form as:

$$\begin{bmatrix} \mathbf{V}_j & -\mathbf{V}_i \end{bmatrix} \begin{bmatrix} \mathbf{h}_i \\ \mathbf{h}_j \end{bmatrix} = \mathbf{0} \ \forall i,j \ i \neq j \tag{3.9}$$

For all pairs (i,j), there are $M(M-1)/2$ different equations of the form of Equation (3.9), which form a linear system of equations from which we want to obtain $\mathbf{h}_i$. For example, for $M=4$, we could group these equations as:

$$\begin{bmatrix} \mathbf{V}_2 & -\mathbf{V}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{V}_3 & \mathbf{0} & -\mathbf{V}_1 & \mathbf{0} \\ \mathbf{V}_4 & \mathbf{0} & \mathbf{0} & -\mathbf{V}_1 \\ \mathbf{0} & \mathbf{V}_3 & -\mathbf{V}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_4 & \mathbf{0} & -\mathbf{V}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_4 & -\mathbf{V}_3 \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \\ \mathbf{h}_4 \end{bmatrix} = \mathbf{0} \tag{3.10}$$

In general, we can rewrite these equations in the form:

$$\mathcal{V}\mathbf{h} = \mathbf{0} \tag{3.11}$$

where

$$\mathcal{V} = \begin{bmatrix} \mathbb{V}_1 \\ \vdots \\ \mathbb{V}_{M-1} \end{bmatrix}, \mathbb{V}_i = \underbrace{\begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}}_{i-1 \text{ blocks}} \underbrace{\begin{bmatrix} \mathbf{V}_{i+1} & \cdots & -\mathbf{V}_i \\ \vdots & \ddots & \vdots \\ \mathbf{V}_M & \cdots & -\mathbf{V}_i \end{bmatrix}}_{M-i+1 \text{ blocks}} \tag{3.12}$$

and

$$\mathbf{h} = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_M \end{bmatrix} \tag{3.13}$$

The matrix $\mathcal{V}$ is of size $(M(M-1)/2)(N_1 - L_2)(N_2 - L_2) \times M(L_1+1)(L_2+1)$ and $\mathbf{h}$ is of size $M(L_1+1)(L_2+1) \times 1$.

In the presence of noise the right hand side of Equation (3.11) is no longer zero and one possible approach is the least-squares solution of Equation (3.11), as:

$$\mathbf{h} = \arg\min \|\mathcal{V}\mathbf{h}\|^2 \tag{3.14}$$

Equation (3.11), or (3.14) in the noisy channel case, states the general principle of the subspace method. The null-vector of $\mathcal{V}$ is the solution $\mathbf{h}$, which by Equation (3.4) is indeed unique if the support is correctly estimated. Therefore, we are interested in characterizing the null-space of $\mathcal{V}$, which can be done through its singular value decomposition (SVD), $\mathcal{V} = \mathbf{A}\mathbf{\Sigma}\mathbf{B}^T$ and analyzing the vectors of $\mathbf{B}^T$ associated with the diagonal elements of $\mathbf{\Sigma}$ which are zero (or close to zero). The $\mathbf{B}^T$ matrix is formed with the eigenvectors of $\mathbf{R}_\mathcal{V} = \mathcal{V}^T\mathcal{V}$, so the solution $\mathbf{h}$ will be the eigenvector with the smallest eigenvalue of $\mathbf{R}_\mathcal{V}$, as proposed as the EVAM algorithm in [62] for 1D and extended to 2D in [65]. Iterative solutions of Equation (3.14) have also been considered for the 1D case [67], with the benefit to easily introduce regularization operators, at the expense of higher computational cost.

3.3.1 Geometrical Constraints

Considering the multiview constraint (2.44), $h_i(\mathbf{x}) = h \circ T_i(\mathbf{x})$, we can reduce the dimensionality of the problem 3.11 while improving its performance against noise and support estimation. Each column in matrix $\mathcal{V}$ represents the values by which each row element of $\mathbf{h}$ is multiplied. The multiview constraint permits to recombine the different columns according to the relationship between the different blurs. The relationship between the different blurs is described by T_i. For example, for two perpendicular filters h_1 and h_2, assuming $L_1 = L_2 = L$, their relationship in two dimensions is simply described as:

$$T_2 = \begin{pmatrix} 0 & -1 & L \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, h_1(\mathbf{x}) = h_2(T_2\mathbf{x}) = h_1(L - y, x) \tag{3.15}$$

and Equation (3.9) can be simplified to:

$$\begin{bmatrix}\mathbf{V}_2 & -\mathbf{V}_1\end{bmatrix}\begin{bmatrix}\mathbf{h}_1 \\ \mathbf{h}_1 \circ T_2\end{bmatrix} = \begin{bmatrix}\mathbf{V}_2 - \mathbf{V}_1 \circ T_2\end{bmatrix}\begin{bmatrix}\mathbf{h}_1\end{bmatrix} = \mathbf{0} \tag{3.16}$$

where $\mathbf{V}_i \circ T_j$ represents a transformation at each row of $\mathbf{V}_i$. The support of the filter h set must contain all non-zero values of all filter h_i.

The same principle applies to any relationship between the different PSF, such as the ones described in Section 2.4 for the multiview case. That is, that the different PSF are rotated versions of a common kernel. We will denote the PSF of the first acquisition, h_1, as the common PSF kernel.

For the example with $M = 4$, Equation (3.11) reduces to:

$$\begin{bmatrix}\mathbf{V}_2 \circ T_2 - \mathbf{V}_1 \circ T_1 \\ \mathbf{V}_3 \circ T_3 - \mathbf{V}_1 \circ T_1 \\ \mathbf{V}_4 \circ T_4 - \mathbf{V}_1 \circ T_1 \\ \mathbf{V}_3 \circ T_3 - \mathbf{V}_2 \circ T_2 \\ \mathbf{V}_4 \circ T_4 - \mathbf{V}_2 \circ T_2 \\ \mathbf{V}_4 \circ T_4 - \mathbf{V}_3 \circ T_3\end{bmatrix}\mathbf{h}_1 = \mathbf{0} \tag{3.17}$$

In the general form:

$$\tilde{\mathcal{V}}\mathbf{h}_1 = \mathbf{0} \tag{3.18}$$

where

$$\tilde{\mathcal{V}} = \begin{bmatrix}\tilde{\mathbb{V}}_1 \\ \vdots \\ \tilde{\mathbb{V}}_{M-1}\end{bmatrix}, \tilde{\mathbb{V}}_i = \begin{bmatrix}\mathbf{V}_{i+1} \circ T_{i+1} - \mathbf{V}_i \circ T_i \\ \vdots \\ \mathbf{V}_M \circ T_M - \mathbf{V}_i \circ T_i\end{bmatrix} \tag{3.19}$$

The matrix $\tilde{\mathcal{V}}$ is of size $(M(M-1)/2)(N_1 - L_2)(N_2 - L_2) \times (L_1+1)(L_2+1)$, and $\tilde{\mathcal{V}}^T\tilde{\mathcal{V}}$ is of size $(L_1+1)(L_2+1) \times (L_1+1)(L_2+1)$ which does not depend on the image size nor the number of channels. Compared to the multichannel case, this represents a reduction by a factor M^2.

Implementation issues

For the general multichannel case, the matrix $\mathcal{V}$ is of size $(M(M-1)/2)(N_1-L_2)(N_2-L_2)\times M(L_1+1)(L_2+1)$, and its direct null-space characterization is typically infeasible. For example, for a two dimensional image of size $(N_1, N_2) = (256, 256)$ pixels, four views ($M = 4$) with blurs of support $(L_1, L_2) = (10, 10)$ pixels and image bit depth of 8 bits, the size of the matrix $\mathcal{V}$ is $363096 \times 484 \approx 168$Mbytes. The matrix $\mathbf{R}_\mathcal{V}$ is of size $M(L_1+1)(L_2+1) \times M(L_1+1)(L_2+1)$, does not depend on the image size, and it is usually a more tractable one. For the aforementioned example, the size of $\mathbf{R}_\mathcal{V}$ is $484 \times 484 \approx 229$kbytes. For the multiview case, matrix $\tilde{\mathcal{V}}$ is of size $(M(M-1)/2)(N_1 - L_2)(N_2 - L_2) \times (L_1+1)(L_2+1)$, that is, for this example, $363096 \times 121 \approx 42$Mbytes. The matrix $\mathbf{R}_{\tilde{\mathcal{V}}}$ is of size $(L_1+1)(L_2+1) \times (L_1+1)(L_2+1)$ which does not depend on the image size nor the number of channels. For this example, $\mathbf{R}_{\tilde{\mathcal{V}}}$ is $121 \times 121 \approx$ 14kbytes, which is a tractable problem. The full eigenvector decomposition of a square

matrix of size (m, m) requires $O(m^3)$ operations. Therefore, by using the multiview constraint the computational complexity is reduced from $O(M^3(L_1+1)^3(L_2+1)^3)$ to $O((L_1+1)^3(L_2+1)^3)$, that is by a factor M^3.

Even in the multiview framework, the computation of the matrix $\mathbf{R}_{\tilde{\mathcal{V}}} = \tilde{\mathcal{V}}^T\tilde{\mathcal{V}}$ by first constructing matrix $\tilde{\mathcal{V}}$ and then perform the matrix multiplication is prohibitive. Matrix $\mathbf{R}_{\tilde{\mathcal{V}}}$ can also be expressed as:

$$\mathbf{R}_{\tilde{\mathcal{V}}} = \sum_{m=1}^{(N_1-L_1)(N_2-L_2)} \tilde{\mathcal{V}}(m, 1..(L_1+1)(L_2+1))^T \tilde{\mathcal{V}}(m, 1..(L_1+1)(L_2+1)) \tag{3.20}$$

where $\tilde{\mathcal{V}}(m, 1..n)$ represents the mth row of $\tilde{\mathcal{V}}$ in Equation (3.18). Therefore, $\mathbf{R}_{\tilde{\mathcal{V}}}$ can be computed directly without the need of the whole $\tilde{\mathcal{V}}$ but only one of its rows at each summatory term. Table 3.1 summarizes the algorithm described in this section. Vect($\bullet$) is the operation of ordering lexicographically an image into vector form, and Unvect($\bullet$) is the inverse operation.

{build $\mathbf{R}_{\tilde{\mathcal{V}}}$}
$\mathbf{R}_{\tilde{\mathcal{V}}} \leftarrow \mathbf{0}$
for all $\mathbf{x} \in (L_1, N_1 - L_1) \times (L_2, N_2 - L_2)$ **do**
  for $i = 1..M-1$ **do**
    for $j = i+1..M$ **do**
      $\tilde{\mathcal{V}} \leftarrow \text{Vect}(v_j \circ T_j(x..x+L_1, y..y+L_2) - v_i \circ T_i(x..x+L_1, y..y+L_2))$
      $\mathbf{R}_{\tilde{\mathcal{V}}} \leftarrow \mathbf{R}_{\tilde{\mathcal{V}}} + \tilde{\mathcal{V}}^T\tilde{\mathcal{V}}$
    end for
  end for
end for
{compute null-space($\mathbf{R}_{\tilde{\mathcal{V}}}$)}
$\mathbf{h}_1 \leftarrow \min \text{eigen}(\mathbf{R}_{\tilde{\mathcal{V}}})$
$h_1 \leftarrow \text{Unvect}(\mathbf{h}_1)$

$h_i \leftarrow h_1 \circ T_i^{-1}$

Table 3.1: Geometrically-constrained PSF estimation algorithm.

3.3.2 Shape Constraints

As shown in the validation of these techniques in Section 3.4, subspace methods have a strong sensitivity to noise and support estimation. However, by further introducing shape constraints, satisfactory results can be obtained. The PSF can be approximated within certain accuracy with a Gaussian shape. This knowledge can be introduced along with the multiview geometrical constraint into the cross-channel relationship. Therefore, the problem is reduced to find the parameters of the introduced model, as:

$$\sigma = \arg\min_{\sigma} \left\| \tilde{\mathcal{V}}\mathbf{h}_1 \right\|^2 \tag{3.21}$$

where the point spread function kernel $\mathbf{h}_1$ is assumed to be Gaussian in all three directions, as:

$$\mathbf{h}_1(\mathbf{x}) = G(0, \sigma_x) \cdot G(0, \sigma_y) \cdot G(0, \sigma_z) \tag{3.22}$$

Equation (3.21) is quadratic with respect to $\mathbf{h}_1$ and can be solved with the golden search technique [51], as detailed in Table 3.2. Again, the matrix product $\tilde{\mathcal{V}}\mathbf{h}$ can be obtained row-by-row, without the need to build the whole $\tilde{\mathcal{V}}$ matrix, as in Section 3.3.1.

$\{\text{find } \sigma = \min f(\sigma) |\ \sigma \in [\sigma_a, \sigma_b]\}$
$f(\sigma) := \tilde{\mathcal{V}}\mathbf{h}(\sigma)$
Require: $\sigma_a < \sigma_c < \sigma_d < \sigma_b$, $f(\sigma_a) > f(\sigma_c)$ and $f(\sigma_d) < f(\sigma_b)$
  repeat
    if $f(\sigma_c) < f(\sigma_d)$ **then**
      $\sigma_b \leftarrow \sigma_d$
      $\sigma_d \leftarrow \sigma_c$
      $\sigma_c \leftarrow \sigma_a + \frac{\sqrt{5}-1}{2}(\sigma_b - \sigma_a)$
    else
      $\sigma_a \leftarrow \sigma_c$
      $\sigma_c \leftarrow \sigma_d$
      $\sigma_d \leftarrow \sigma_b - \frac{\sqrt{5}-1}{2}(\sigma_b - \sigma_a)$
    end if
  until $(f(\sigma_c) - f(\sigma_d)) < \epsilon$

  $\sigma = \min f(\sigma) \approx \sigma_c \approx \sigma_d$

Table 3.2: Shape-constrained PSF estimation algorithm

3.4 Validation

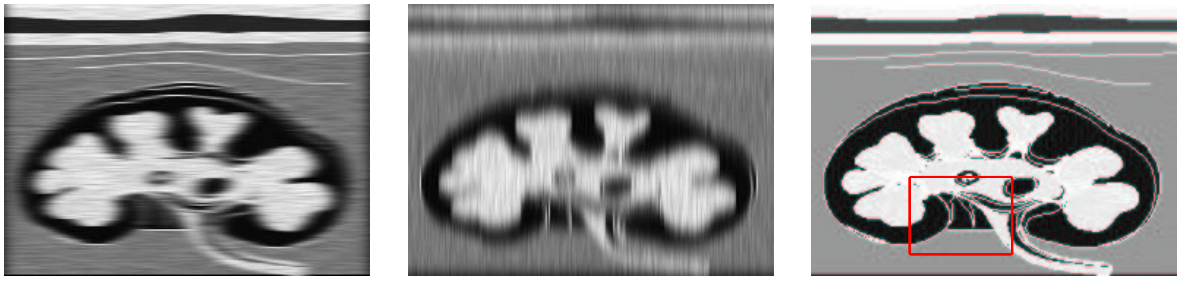
In this section we will evaluate the techniques of point spread function estimation presented in this chapter, and the deconvolution techniques presented in Chapter 2. We will use synthetic images for their evaluation, and application to real ultrasound medical images will be done in Chapters 5 and 6.

3.4.1 Methods

Synthetic Images

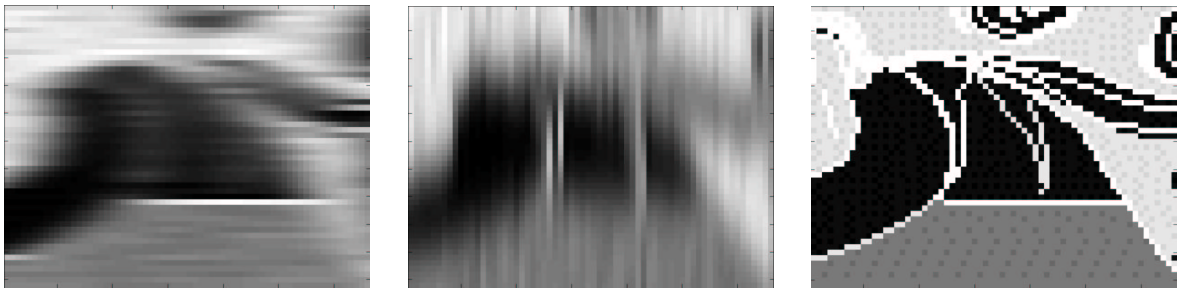
For the synthetic images, we have artificially added noise and blurred a 2D synthetic image of a kidney [78]. In order to mimic log-compressed ultrasound speckle noise characteristics, we have added speckle noise such as $v_{noisy} = v + n \cdot \sqrt{v}$, where n is uniformly distributed random noise with zero mean and variance σ_n^2, and v the original intensity value [88]. The noisy image was blurred with an oriented Gaussian kernel with a variance σ_k pixels at 0 and 90 degrees, following the colored Gaussian

speckle model for ultrasound log-compressed images presented by Kao *et al.* [79] (see Appendix A.3.2 for details). The used values for σ_n and σ_k were $\{0.005, 0.05, 0.5\}$ and $\{2, 5, 8\}$ respectively. Other authors log-decompress the image before processing it, although this presents two main issues: level quantization and knowing the parameters involved in the log-compression (Appendix A). Figure 3.1 shows the example images used to present the techniques, which correspond to parameters $\sigma_n = 0.05$ and $\sigma_k = 5$. Experimentation with other parameters will be shown latter, when comparing the methods. Figures 3.1.a and 3.1.b are the simulated views at two perpendicular orientations, with uncorrelated speckle added noise. We will denote these simulated views as v_0 and v_{90}. Figure 3.1.c is the original image, which we will denote as v. Notice that horizontal details are kept and vertical details are lost in v_0, and the other way around for v_{90}. This is better appreciated on the zoom on a region of interest (ROI) in Figure 3.2. The goal is to fuse information from v_0 (Figure 3.1.a) and v_{90} (Figure 3.1.b) to recover the original image v (Figure 3.1.c).



(a) (b) (c)

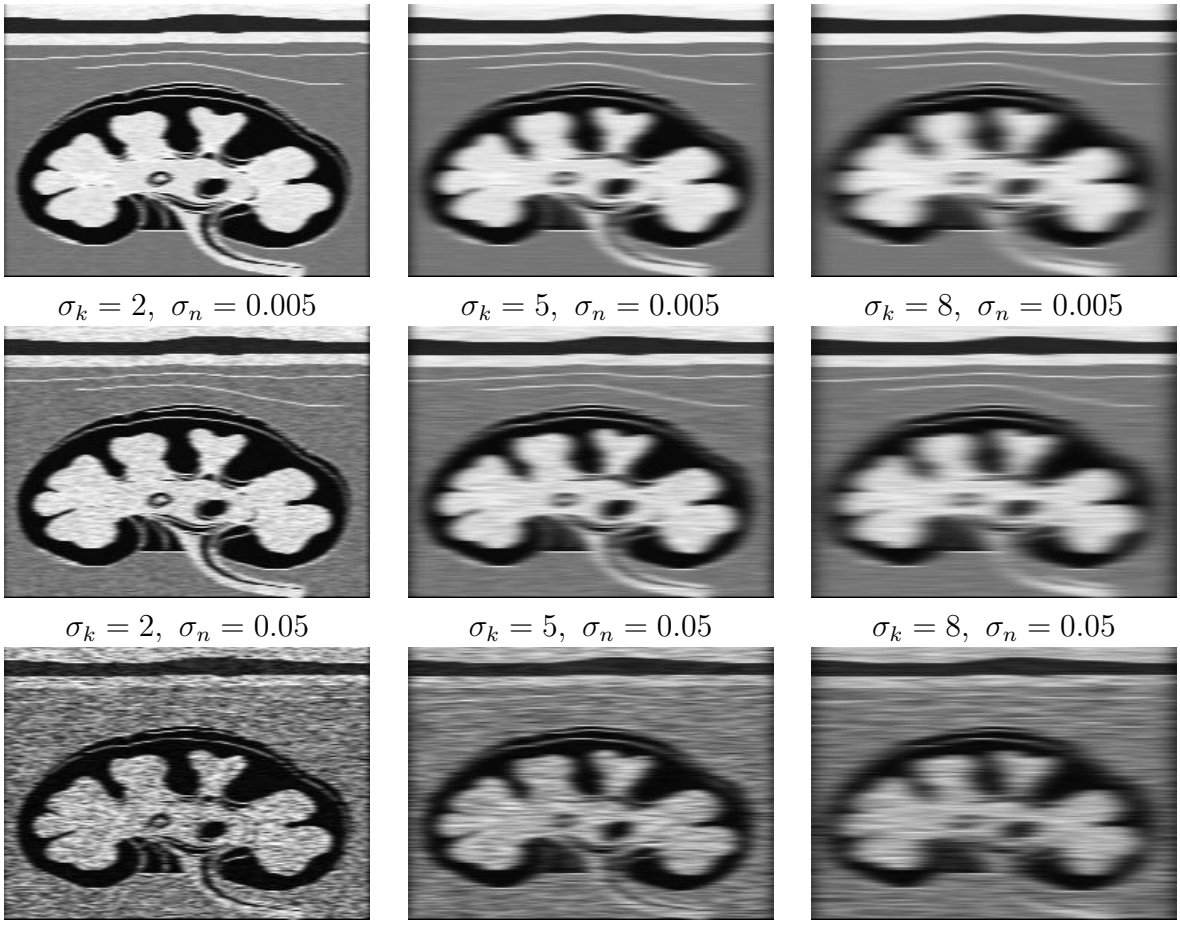
Figure 3.1: Synthetic images for noise variance $\sigma_n^2 = 0.005$ and blur variance $\sigma_k = 5$. a) v_0, noised and blurred image at 0 degrees (35.28 dB), b) v_{90}, noised and blurred image at 90 degrees (29.06 dB), c) Original image (∞ dB), with box indicating a region of interest (ROI). The goal is to recover image c) from a) and b).



(a) (b) (c)

Figure 3.2: Zoom on a region of interest (ROI) of synthetic images. a) v_0, noised and blurred image at 0 degrees, b) v_{90}, noised and blurred image at 90 degrees, c) original image.

Figures 3.3 and 3.4 show the simulated v_0 and v_{90} respectivley, for different conditions of noise and blur variance.



$\sigma_k = 2,\ \sigma_n = 0.005$ $\sigma_k = 5,\ \sigma_n = 0.005$ $\sigma_k = 8,\ \sigma_n = 0.005$

$\sigma_k = 2,\ \sigma_n = 0.05$ $\sigma_k = 5,\ \sigma_n = 0.05$ $\sigma_k = 8,\ \sigma_n = 0.05$

$\sigma_k = 2,\ \sigma_n = 0.5$ $\sigma_k = 5,\ \sigma_n = 0.5$ $\sigma_k = 8,\ \sigma_n = 0.5$

Figure 3.3: Simulated acquisition v_0 for different levels of blur (σ_k) and noise (σ_n).

Metrics

It is a difficult task to quantify the performance of the different algorithms in a meaningful way. To measure the difference between two images, one of most used measures is the peak signal-to-noise ratio (PSNR) , which requires the knowledge of the original image to be computed. The PSNR between the images $A(\mathbf{x})$ and $B(\mathbf{x})$, measured in dB, is defined as:

$$PSNR = 20 \log_{10} \left(\frac{\max A(\mathbf{x})}{RMSE} \right) \tag{3.23}$$

where RMSE is the root mean squared error. It is defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum \|A(\mathbf{x}) - B(\mathbf{x})\|^2} \tag{3.24}$$

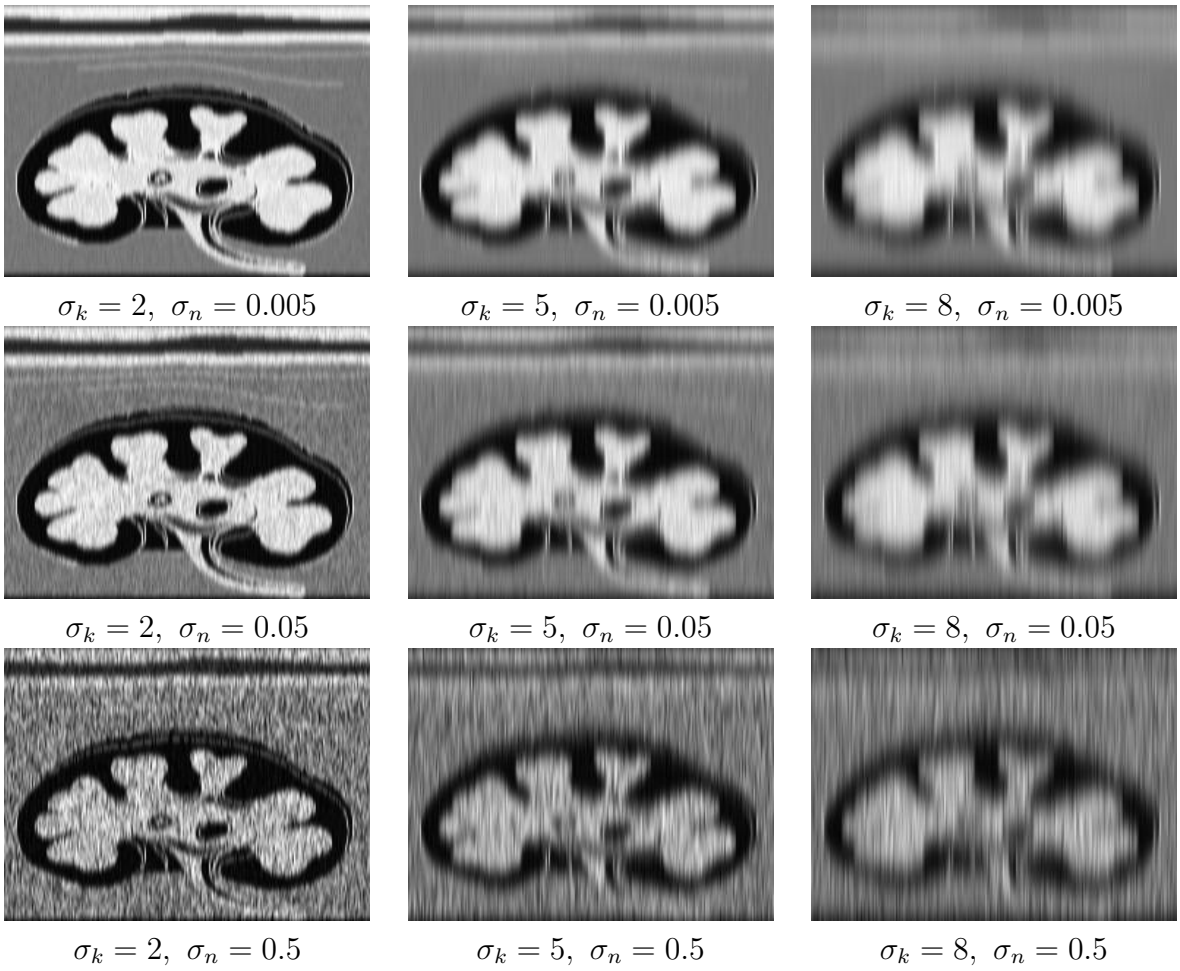


Figure 3.4: Simulated acquisition v_{90} for different levels of blur (σ_k) and noise (σ_n).

where n is the total number of pixels in the overlapping area between images $A(\mathbf{x})$ and $B(\mathbf{x})$.

Other metrics are found in the literature, based on evaluating the focusing of the image, such as the ones based on the gray-level variance [146], the energy of Laplacian [147], wavelet focusing [84] or lateral correlation [160]. These metrics do not need the original image to be computed.

Visual evaluation also plays a critical role on comparing image processing techniques, since many parameters of interest are hardly described using these quantitative measures. In some cases, specific measures can be defined for the parameters of interest of each application. For example, tissue delineation can be measured by the relative contrast between two tissues.

3.4.2 PSF estimation

We evaluated on the synthetic images of Figure 3.1 the presented techniques to estimate the point spread function. These techniques are: unconstrained subspace tech-

niques (denoted as *ss*), constrained by the multiview geometrical relationship (denoted as *mv*), and further constrained by a Gaussian shape (denoted as *gmv*).

Table 3.3 summarizes the results of this experiment. As it can be observed, in the noiseless case all techniques successfully retrieve the original synthetic shape of the blur, up to a great level of accuracy. However, even for the smallest level of noise, the unconstrained subspace techniques fail to estimate the size of the blur. These cases (indicated with "x" in Table 3.3), fail in retrieving any coherent shape for the PSF, creating erratic results. This reveals their strong sensitivity on noise.

		$\bar{\sigma_k}$		
σ_k	σ_n	*ss*	*mv*	*gmv*
2	0	2.00	2.00	1.97
5		5.00	5.00	5.13
8		8.01	8.01	8.14
2	0.005	x	2.14	2.05
5		x	4.75	5.09
8		x	8.34	8.17
2	0.05	x	x	2.21
5		x	x	5.07
8		x	x	8.54
2	0.5	x	x	2.15
5		x	x	5.32
8		x	x	9.12

Table 3.3: Point spread function variance estimation ($\bar{\sigma_k}$) for different levels of noise σ_n and sizes of blur σ_k. The different techniques are: subspace techniques (*ss*), using the multiview constraint (*mv*), and the multiview constraint and the Gaussian shape *a priori* knowledge (*gmv*). Experiments marked with "x" indicate that the algorithm failed.

Introducing the multiview constraint helps improving this limit, accepting a small quantity of noise. However, it also fails for a certain level of noise, which limits its applicability on real data. Introducing the constraint of Gaussian shape of the kernel, the technique successfully estimates the right size of the blur at all tested levels of noise.

The subspace and the multiview techniques need an estimation of the size of the support. The results reported here correspond to the best estimation of this size, done by exhaustive search. Although there are some techniques to optimize this search, it is not obvious to certify that the best result is obtained. When shape constraints are introduced, there is no longer the need of estimating the support of the PSF since it is implicitly assumed with the chosen model.

3.4.3 Image Restoration

Using the multiview PSF estimation technique with Gaussian shape constraint, we evaluate the deconvolution techniques presented in Chapter 2 on these synthetic images. Table 3.4 describes the results of this experiment using three different regularization potential functions, namely: Tikhonov, Huber and Lorentz. As it can be observed, the latter performs best in all cases for different levels of regularization. Quantitatively, the non-regularized scheme ($\lambda = 0$) has a slightly lower but comparable performance to the Lorentz regularization. For low and medium levels of noise, the Tikhonov and the Huber regularization functions perform worse than the Lorentz functions. However, for high-levels of noise Tikhonov and Huber functions slightly outperform Lorentz function. This is due to the fact that the Lorentz potential function preserves most edges, but also noise patterns at a certain level of noise power.

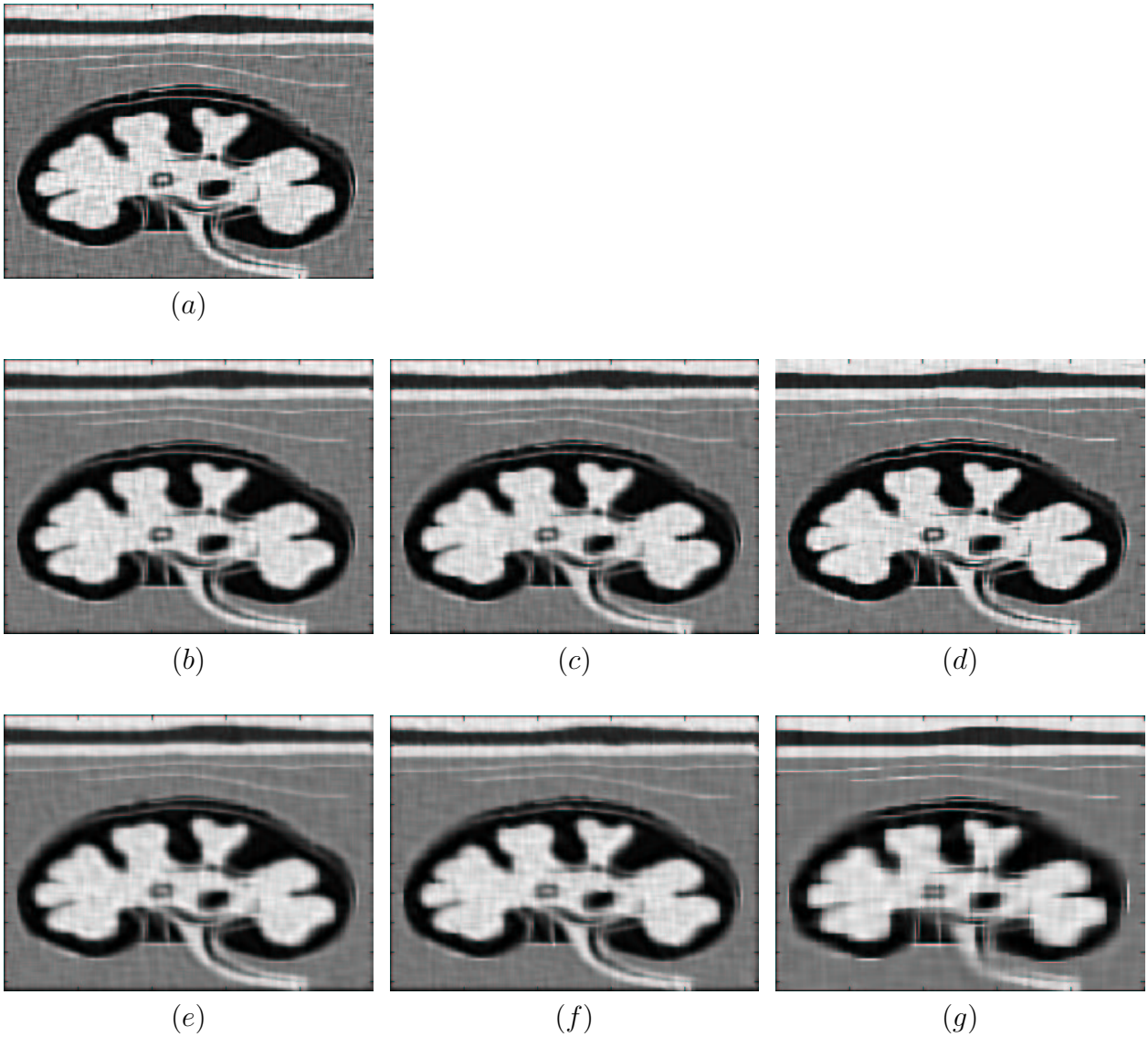
The influence of the regularization parameter λ shows that an over-regularized solution decreases the fidelity of the reconstruction. For Tikhonov and Huber potential functions, better results are obtained with balanced trade-off of the data fidelity term and the regularization term. For Lorentz potential function, results are quantitatively similar, which might be an indication that this potential function is appropriate to describe prior knowledge of the original image. Although this fact might be justified by the synthetic nature of the images, such behavior is adequate also for real images.

PSNR				-	Tikhonov		Huber		Lorentz	
σ_k	σ_n	v_0	v_{90}	$\lambda = 0$	$\lambda = 0.5$	$\lambda = 1$	$\lambda = 0.5$	$\lambda = 1$	$\lambda = 0.5$	$\lambda = 1$
2	0.005	41.40	35.51	50.64	42.20	38.89	42.97	39.33	52.11	52.16
5		36.10	29.63	41.98	37.92	35.72	38.08	35.85	42.41	42.33
8		33.14	26.60	38.92	35.96	34.03	36.05	34.04	39.16	39.10
2	0.05	39.72	34.78	43.98	40.33	37.78	40.72	38.29	45.61	46.61
5		35.28	29.06	39.79	36.98	35.02	37.07	35.20	40.59	40.67
8		32.57	26.25	37.71	35.28	33.55	35.39	33.55	38.06	38.15
2	0.5	30.36	28.06	26.92	30.77	30.53	30.77	30.67	27.27	27.35
5		29.56	25.87	29.25	30.48	29.94	30.54	30.09	29.57	29.46
8		28.20	24.02	29.69	30.05	29.34	29.90	29.26	30.01	30.14

Table 3.4: Peak Signal to Noise Ratio (dB) for blind multiview deconvolution of simulation data, for degradations σ_k and speckle noise variance σ_n. Regularization methods: none(-), Tikhonov, Huber and Lorentz, for values of $\lambda = 0.5$ and $\lambda = 1$.

Figure 3.5 shows the deconvolution results on the synthetic images for medium levels of noise ($\sigma_n = 0.05$) and blur ($\sigma_k = 5$). A zoom on a region of interest can be observed in Figure 3.6. It can be visually appreciated that the non-regularized solution shows amplified noise. This puts in evidence that simple metrics as PSNR are not sufficient to describe the performance of these algorithms. Indeed, the non-regularized technique minimizes explicitly the squared differences with the simulated acquisitions, therefore it makes sense that numerically it obtains satisfactory results.

The effect of regularization can be further appreciated on the zoomed images. While Tikhonov blurs uniformly all tissue boundaries, Huber preserves some of those, although it created a bipolar behavior due to the abrupt change in its derivative. Edges that are over the threshold α are preserved and those below not. Lorentz function offers a smoother behavior than other techniques, preserving most of the visible edges present in the original acquisitions.



(a) (b) (c) (d) (e) (f) (g)

Figure 3.5: Blind multiview deconvolution for $\sigma_k = 5$ and $\sigma_n = 0.05$, using different regularization potential functions: top, a) no regularization ($\lambda = 0$). Middle, $\lambda = 0.5$, b) Tikhonov, c) Huber, d) Lorentz. Bottom, $\lambda = 1$, e) Tikhonov, f) Huber, g) Lorentz.

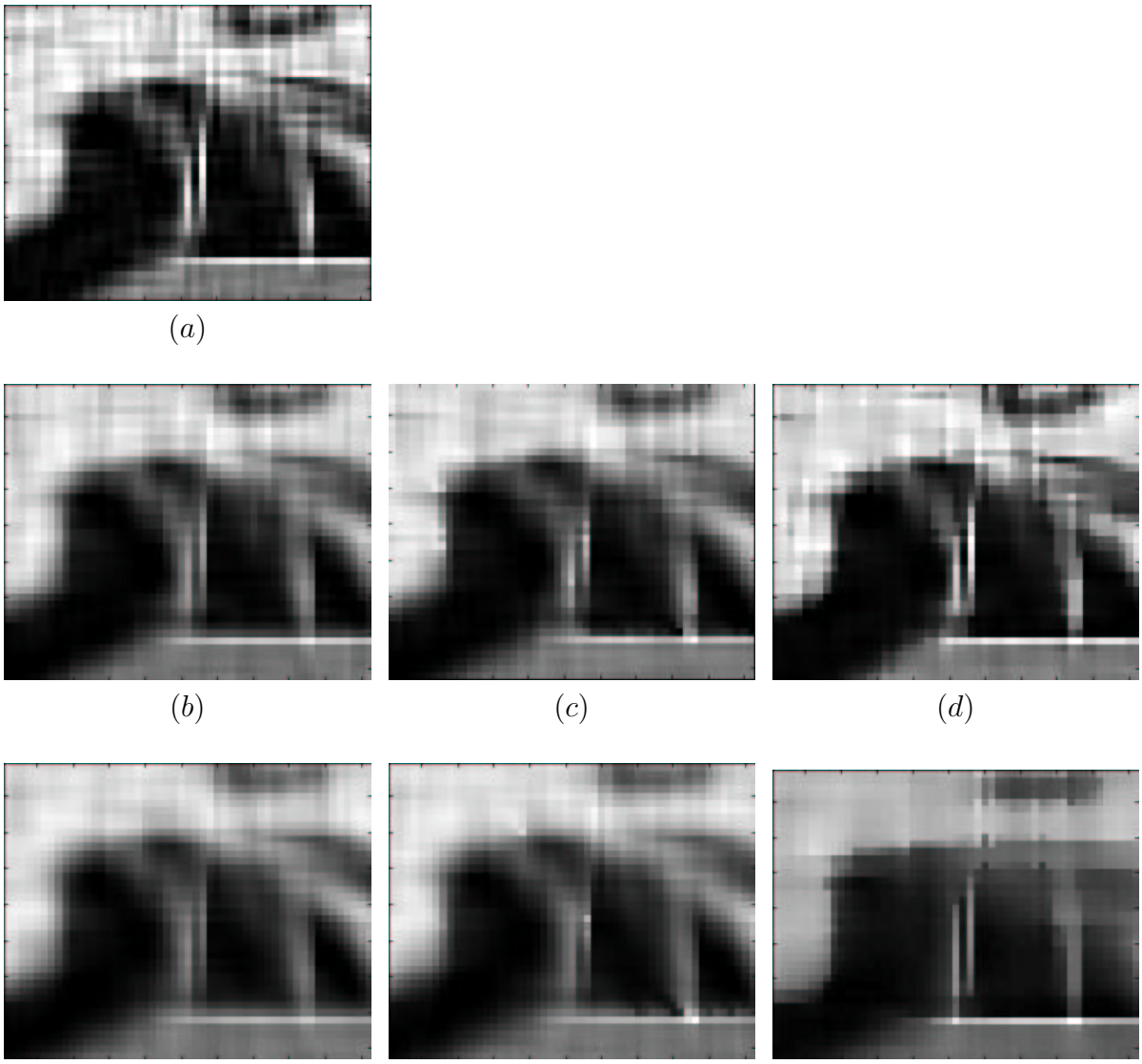


Figure 3.6: Region of interest for blind multiview deconvolution for $\sigma_k = 5$ and $\sigma_n = 0.05$, using different regularization potential functions: top, a) no regularization ($\lambda = 0$). Middle, $\lambda = 0.5$, b) Tikhonov, c) Huber, d) Lorentz. Bottom, $\lambda = 1$, e) Tikhonov, f) Huber, g) Lorentz.

3.5 Conclusions

In this chapter we developed a technique to estimate the point spread function from anisotropically degraded acquisitions. We first reviewed the blind deconvolution techniques for the single channel case, which are based solely on the intrinsic differences between the implicit original image and the point spread function. Afterward, we reviewed the multichannel techniques, which use the diversity of the different channels to estimate the different point spread functions.

In particular, we focused on the subspace techniques, which offer a closed-form solution of the PSF via the computation of the null-space of a matrix constructed from the cross-channel relationship. However, this technique is quite sensitive to noise and to the correct estimation of the support of the PSF. We then introduced the geometrical relationship between the different PSF, which reduces the dimensionality of the problem, thus improving the robustness against noise and reducing the computational time. However, these improvements may not be sufficient for real ultrasound medical imaging. Then, we further introduced shape constraints to the PSF kernel, in particular, a Gaussian shape. This constrained system succeeded to estimate within reasonble tolerance the appropriate parameters of the system point spread function from the different acquisitions.

We validated these techniques on synthetic data. The results obtained with the shape constrained technique were used along the deconvolution methods presented in Chapter 2 to restore the synthetic images. We tested three different regularization functions, namely: Tikhonov, Huber and Lorentz. The latter performed best in all cases with respect to the signa-and-noise ratio. Application to real ultrasound data is left for Chapters 5 and 6.

CHAPTER 4

Multiview Fusion

4.1 Introduction

In Chapters 2 and 3, we have described the multiview deconvolution technique, which restores the original volume by estimating the degradation process of the system. In this chapter, we present a set of different techniques, based on a fusion approach, consisting in building the reconstructed volume by capturing the valuable information from each view. The goal is to preserve salient features and the best spatial resolution, while reducing speckle patterning at the same time.

There are mainly two benefits of using such fusion approaches in comparison to deconvolution: speed and simplicity. The presented techniques are computed in a single step, which reduces the computational cost with respect to iterative methods. Convergence is guaranteed by construction, since the reconstructed image is a combination of the input images. There is no need for *a priori* information about the system, such as the point spread function or the angle of acquisition, leading to better robustness with respect to the estimation of these parameters. However, by not using this information, the maximum performance of these methods is lower than the one of deconvolution approaches. Indeed, these techniques just preserve and combine the information from each view, but do not really reverse the degradation process.

In order to illustrate and evaluate the methods both quantitatively and qualitatively, we used the synthetic images presented in Section 3.4. Application to real data will be presented in Chapters 5 and 6. Registration of the different views is assumed to be perfect throughout the chapter. The actual registration issues and the techniques to overcome them in different applications will also be discussed in each corresponding chapter.

We divide the set of approaches into two classes, depending on the domain in which they operate. Methods operating in the spatial domain are described in Section 4.2, while those working in the spectral domain are presented in Section 4.3. A comparison is performed in Section 4.4, and finally, conclusions are drawn in Section 4.5.

4.2 Spatial Domain Fusion

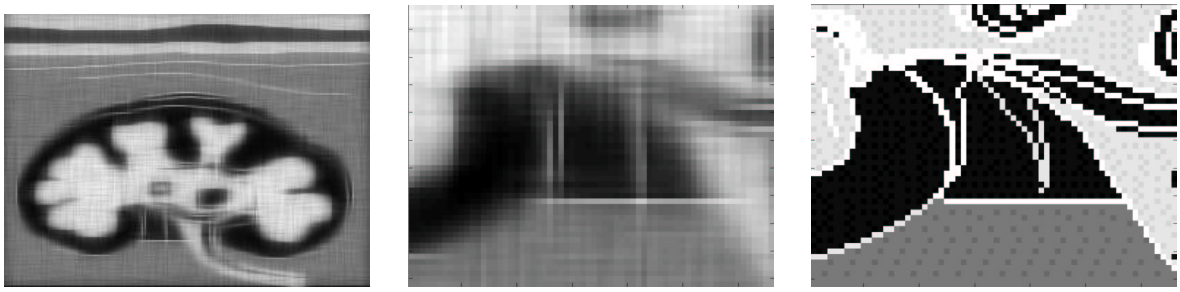
4.2.1 Averaging

Averaging is commonly used for spatial compounding, due to the ability to reduce uncorrelated noise from the different views [155]. Moreover, it is robust and fast. However, the resulting images do not preserve the best spatial resolution and do not guarantee feature preservation, that is, features present in only one of the views might be lost. Denoting by v_i each of the views, the estimated volume $\tilde{v}$ is defined by:

$$\tilde{v}(\mathbf{x}) = \frac{1}{M} \sum_{i=1}^{M} v_i(\mathbf{x}) \tag{4.1}$$

where M is the number of views, and $\mathbf{x} = [x\ y\ z]^T$ is a point in space.

Figure 4.1 shows the averaging of the example images, along with a zoom on the ROI defined in Figure 3.2. Both vertical and horizontal features are visible, but with a poor contrast and tissue delineation severely affected by blurring. The difference with the original image, measured as the PSNR (see Section 4.4) is 33.32 dB, which is in between the PSNR of each individual view v_0 and v_{90} (35.28 dB and 29.06 dB respectively). In other words, quantitatively the reconstructed image differs more than the original image than from the single view v_0. This can be explained by the fact that the original image has a high horizontal content, better preserved in v_0 than in v_{90}.



(a) (b) (c)

Figure 4.1: a) Average image (PSNR = 33.32 dB), b) ROI on average image, c) ROI on original image

4.2.2 Weighted averages

One approach to overcome the loss of features when performing averaging is to give different weights to each view, depending on the presence of specific features, or its *saliency*. We denote weighted averaging as:

$$\tilde{v}(\mathbf{x}) = \sum_{i=1}^{M} \alpha_i(\mathbf{x}) v_i(\mathbf{x}) \tag{4.2}$$

where $\alpha_i(\mathbf{x})$ represents the weights of each measurement. These weights are positive, and sum up to unity ($0 \leq \alpha_i \leq 1$, $\sum_i \alpha_i = 1$). Simple averaging is obtained for $\alpha_i = \frac{1}{M}$.

Several saliency measures have been proposed in the literature. The measures are based on local statistics [33], such as local contrast, or use some segmentation algorithms, such as a Canny edge detector (see a review in [123]). For spatial compounding, Leotta *et al.*[100] proposed to use the local incidence angle of the ultrasound beam. For the three-dimensional breast imaging case, Krucker *et al.*[90] showed qualitative better results using local contrast as saliency measure. Depending on the size of the neighborhood of local statistics, finner or larger details will be taken into account into the saliency map. This has lead also to some multiscale approaches, which combine saliency at different levels of detail permitting the detection of features at the desired scale.

In our experiments, we used local contrast as saliency our measure:

$$sal_i(\mathbf{x}) = \left| v_i(\mathbf{x}) - \bar{v_i}(\mathbf{x}) \right| \tag{4.3}$$

where $\bar{v}(\mathbf{x}) = \frac{1}{N} \sum_{j=1}^{N} v(\mathbf{x}_j)$ is the local intensity mean over N neighbors.

When the size of a feature is close to the size of a speckle pattern, they are hard to distinguish. In order to discriminate specular reflectors from speckle patterning, noise models for speckle have been studied in the literature (e.g. on envelope image [159, 35, 124], on log-compressed images [45, 80, 79, 9]). On the envelope image, speckle shows a Rician distribution [159], which tends to a Rayleigh distribution for low SNR conditions. These distribution are completely characterized by their first two moments, mean and standard deviation. The coefficient of variation (CV) is defined as the ratio:

$$CV = \frac{\sigma_n}{\bar{n}} \tag{4.4}$$

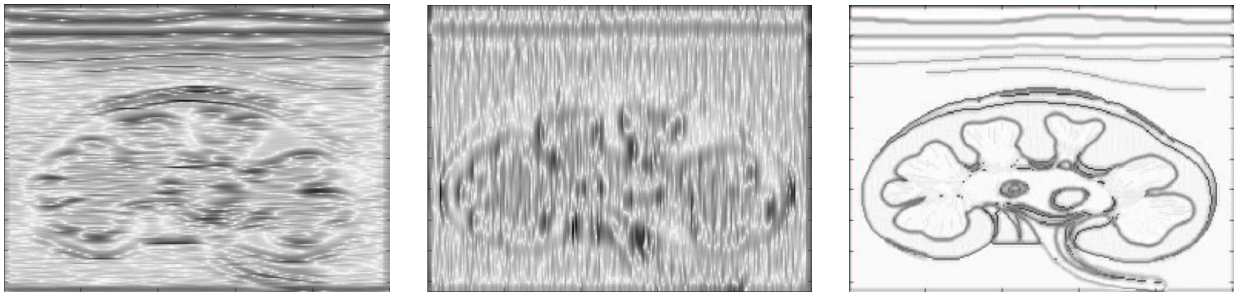
where σ_n is the noise standard deviation and $\bar{n}$ is the noise mean on a local neighborhood. The CV theoretically remains constant in uniform areas and has been widely used in coherent imaging to discriminate speckle patterning, such as in the Lee [97] and Frost [52] filters. A Rayleigh distribution satisfies $CV^{-1} = 1.91$, which is a typically verified in envelope images with low SNR. However, in log-compressed B-mode ultrasonic images, speckle theoretically follows a double exponential distribution [45], although in can be approximated by a colored Gaussian noise [79] of variance independent of the mean. Therefore, as a speckle measure (spk) for log-compressed images we may use, rather than the CV, the ratio of the local standard deviation from noise as:

$$spk_i(\mathbf{x}) = \frac{\sigma_n}{\sigma_{v_i(\mathbf{x})}} \tag{4.5}$$

where the parameter σ_n^2 was estimated in flat areas of the image, where the variance of the image was minimal. Values over unity are clipped, as to have a measure which takes values from one, when a pixel belongs to a speckle area, to zero, when the pixel does not correspond to a speckle area. Therefore, the measure $(1 - spk_i)$ measures the non-speckleicity.

We combined saliency and speckle measures as $\alpha_i^{'} = sal_i * (1 - spk_i)$, and then renormalized as $\alpha_i = \frac{\alpha_i^{'}}{\sum_j \alpha_j^{'}}$.

Figure 4.2 shows the saliency maps for the synthetic image examples. Figure 4.3 shows the actual weights $\alpha(\mathbf{x})$, which combine the saliency and speckle measures. Figure 4.4 shows the resulting image that, compared to uniform averaging, enhances the contrast of features.



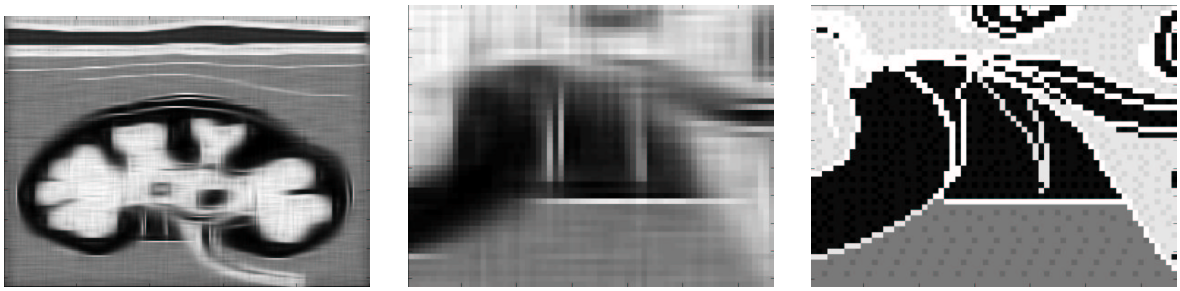
(a) (b) (c)

Figure 4.2: Saliency for synthetic example images (for clarity, black indicates high contrast). a) v_0, b) v_{90}, c) original image.

(a) (b) (c)

Figure 4.3: Saliency combined with speckle detection for synthetic example images (for clarity, black indicates high weight). a) v_0, b) v_{90}, c) original image.

The main problems of weighted averaging are the difficulty of feature identification and the inability to maintain spatial resolution. The first is due to the fact that speckle patterning and features are not easily distinguished by local statistics. As mentioned, multiscale approaches have been suggested to overcome these problems [76], implying new parameters which have to be tuned for each imaging application. Further analysis of the image to find the internal structure of the image, e.g. via the Hessian of the image [9, 161], has been suggested. The problem of feature detection in the multiview case is simpler than in a single image, since the task is to recognize which of the views shows an stronger feature, rather than selecting features in an absolute sense. Even then, it is still not robust, due to the intrinsic similarity between features and speckle



(a) (b) (c)

Figure 4.4: a) Weighted average image (34.02 dB), b) ROI on weighted average image, c) ROI on original image.

patterning. In second place, the effects of different point spread functions for each view are difficult to be taken into consideration within this framework.

4.2.3 Generalized averages

The presented averaging methods tend to blur the images. In fact, it would be desirable to use different operators, function of the data characteristics in each view. Figure 4.5 shows the three main cases involved in image fusion in multiview echography with simple pixel-based operators. In the case of pure speckle patterning, averaging between the uncorrelated views will lead to the desired mean value [155]. In order to keep echogenic (white) features, the maximum operator would be appropriate. To keep the best spatial resolution, the minimum operator should be used, in order to minimize the effects of the point spread function. For hypoechogenic (dark) features, the functions of minimum and maximum should be interchanged. We will not discuss this case, since we are primarily interested in echogenic features. Figure 4.6 shows the results of using the minimum, average and maximum operators to all the example images.

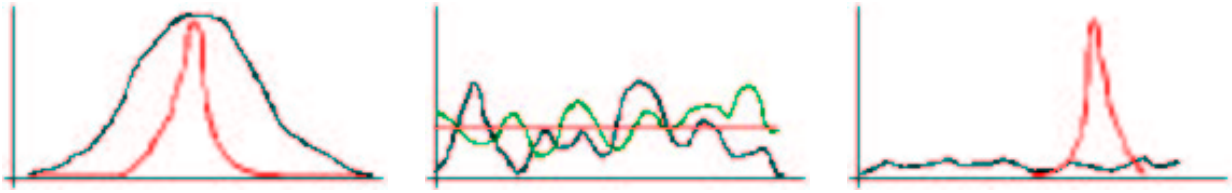
Generalized averaging provides a way to switch from minimum to average and to maximum operators in a smooth fashion [30]. In its general form, it can be described as:

$$\tilde{v}(\mathbf{x}) = f^{-1}\left(\frac{1}{M}\sum_{i=1}^{M} f\left(v_i(\mathbf{x})\right)\right) \tag{4.6}$$

where f is a monotonous function. For $f(\lambda) = \lambda^{\beta}$, generalized averages take the following form:

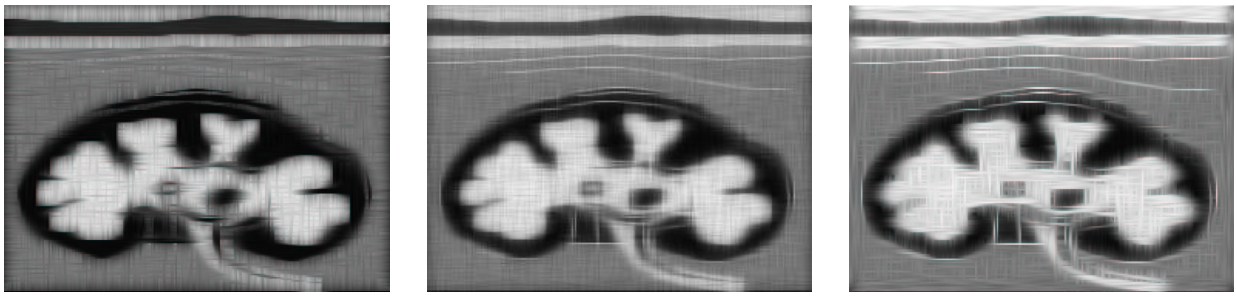
$$\tilde{v}(\mathbf{x}) = \left(\frac{1}{M}\sum_{i=1}^{M} v_i(\mathbf{x})^{\beta(\mathbf{x})}\right)^{1/\beta(\mathbf{x})} \tag{4.7}$$

For $\beta = 1$, an arithmetic average is obtained. For $\beta = 2$, a quadratic average is



(a) (b) (c)

Figure 4.5: Data types in multiview echography and desired operation: a) data agreement, but different PSF (min), b) speckle (average), c) data disagreement, features present in only one view (max).



(a) (b) (c)

Figure 4.6: Example of min, average and max operators on synthetic images, a) min (no blur), b) average (denoising), c) max (feature preservation).

obtained. For $\beta \to \infty$, the generalized average behaves as a maximum operator, and for $\beta \to -\infty$ it behaves as a minimum.

The value of $\beta(\mathbf{x})$ should depend on the data "agreement" between the different views rather than on the saliency of each view. Notice that, in comparison to the weight maps, there is only one β map and not one for each view. The "agreement", or its contrary, "discrepancy" (dis), is measured as the variance of gray levels on a local neighborhood against the mean compounded volume, as:

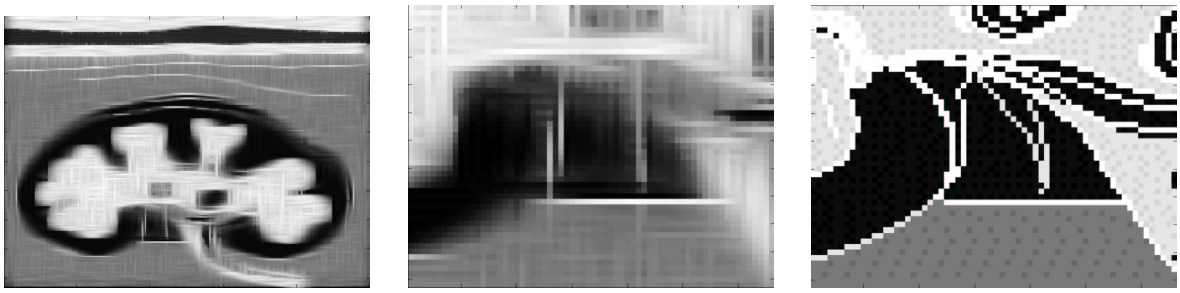
$$dis(\mathbf{x}) = \sum_i \left(v_i(\mathbf{x}) - \sum_i \bar{v_i}(\mathbf{x}) \right)^2 \tag{4.8}$$

The final $\beta(\mathbf{x})$ parameter is computed, after normalization of parameters disagreement (dis) and speckle (spk) as:

$$\beta(\mathbf{x}) = dis(\mathbf{x}) * (1 - \frac{1}{M} \sum_i spk_i(\mathbf{x})). \tag{4.9}$$

In our experiments, in order to accommodate these values to the desired range, a look-up table was used, which mapped the $[0, 1]$ range to $[-100, 100]$. The upper (lower) limit was chosen to have a behavior close (more than 99%) to the maximum (minimum) operators (ideally it would be ∞).

Figure 4.7 shows the results obtained with this technique on synthetic images. It can be observed that edges are slightly better preserved than in the previous techniques. However, as with weighted averages, discrepancy maps are not robust, due to the difficulty to differentiate speckle from features.

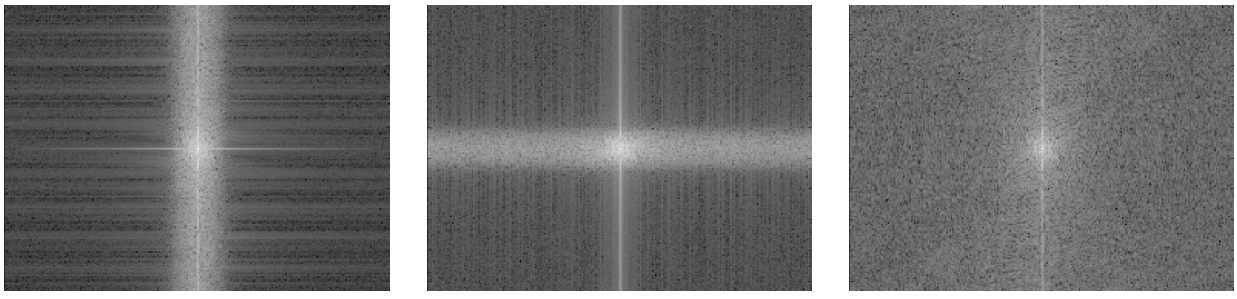


(a) (b) (c)

Figure 4.7: a) Generalized average image (35.21 dB), b) ROI, c) ROI on original image.

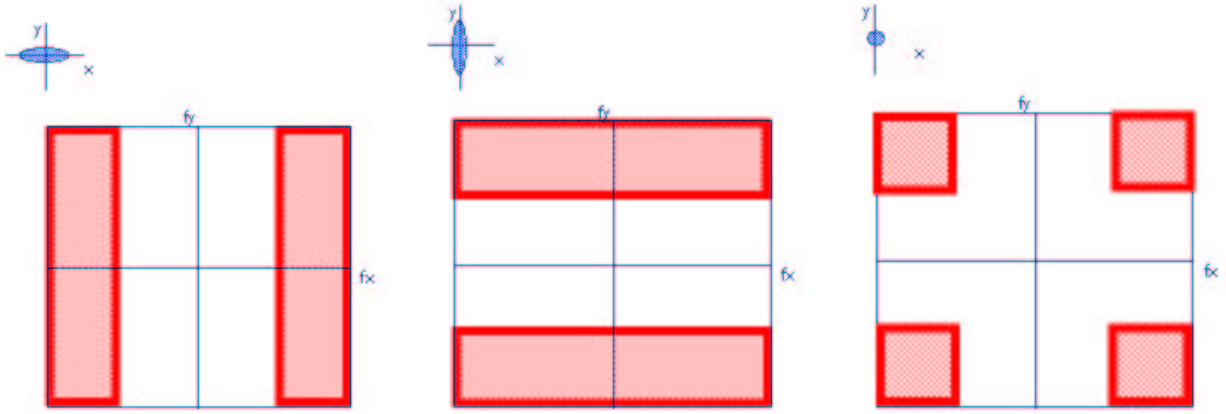
4.3 Spectral Domain Fusion

Another possible approach to preserve the best spatial resolution is to combine the different views in the spectral domain. Depending on the PSF of each view, we have a different spectral coverage (Section 2.4.2), as shown in Figure 4.8 for the synthetic example images. It can be appreciated how v_0 preserves the vertical frequencies, while v_{90} preserves the horizontal ones. The principle of frequency fusion consists in preserving all spectral content, that is keeping in the fused volume the largest components of each view. This principle is schematized in Figure 4.9.



(a) (b) (c)

Figure 4.8: Actual frequency content of example synthetic images (amplitude in logarithmic scale). a) $V_0(\mathbf{f})$, b) $V_{90}(\mathbf{f})$, c) $V(\mathbf{f})$ (original).



(a) (b) (c)

Figure 4.9: Point spread function and associated spectral coverage (plain white area). a) horizontal PSF, vertical frequencies are preserved b) vertical PSF, horizontal frequencies are preserved, c) combined spectral coverage, which should include the contents of the different views.

4.3.1 Fourier Transform Fusion

A straightforward method to combine the spectral content of the different views is by combining their Fourier Transform coefficients. Since we are interested in not using *a priori* information on the PSF, a simple approach to combine these acquisitions is by selecting the coefficient of maximum amplitude at each frequency. Denoting the Fourier transforms of $v_i(\mathbf{x})$ by $V_i(\mathbf{f})$ the fused volume can be written as:

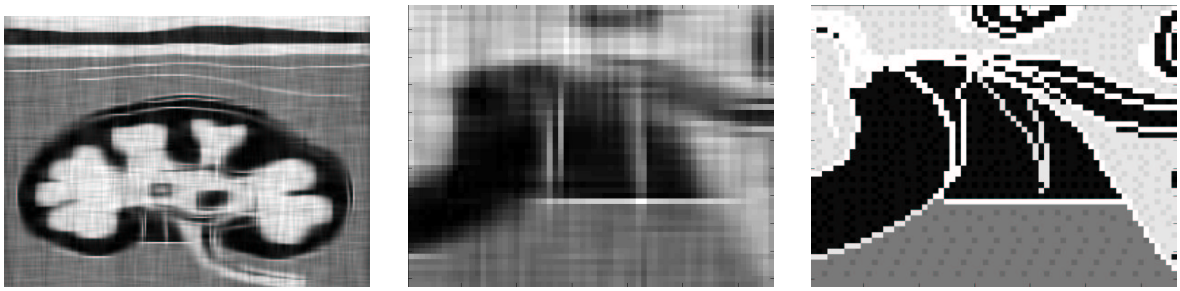
$$\tilde{V}(\mathbf{f}) = max_i \left\{ V_i(\mathbf{f}) \right\} \tag{4.10}$$

where $\max_i \{\bullet\}$ refers to selecting the Fourier component (real and imaginary parts) with the highest absolute value ($|V(\mathbf{f})|$). The actual image is obtained by taking the real part of the inverse Fourier Transform, as:

$$\tilde{v}(\mathbf{x}) = \Re \left[\mathcal{F}^{-1} \left\{ \tilde{V}(\mathbf{f}) \right\} \right] \tag{4.11}$$

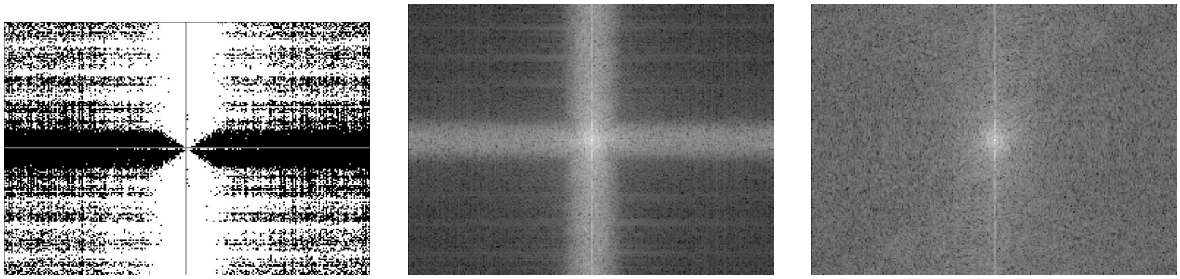
While keeping all spectral information, this method does not guarantee the preservation of features nor the reduction of speckle. On the contrary, noise is preserved at all frequencies. Indeed, depending on the images, the steep changes introduced by the maximum operator in the low frequency coefficients may cause low-frequency artifacts, and similarly, the abrupt changes in high frequency coefficients may cause ringing artifacts. Indeed, changes in one coefficient may have a global effect, and propagate to the whole image in the spatial domain.

Figure 4.10 shows an example of result with this technique. Figure 4.11.a shows from which view each frequency is taken, along with the final frequential content (Figure 4.11.b). Notice how it differs from the frequential content of the original image (Figure 4.11.c).



(a) (b) (c)

Figure 4.10: a) Maximum Frequency fusion (37.31 dB), b) ROI, c) ROI on original image.



(a) (b) (c)

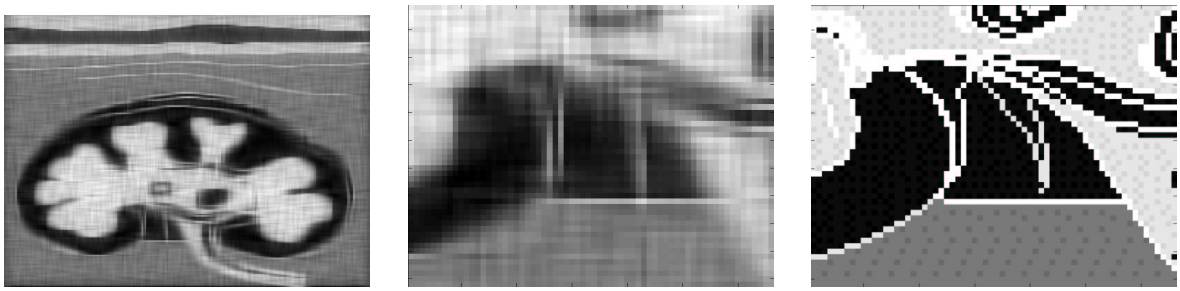
Figure 4.11: a) Frequency mask, white corresponds to v_0, black corresponds to v_{90} , b) Frequency content of the maximum frequency fusion, c) Frequency content of the original image.

One possible performance improvement of this method is to differently treat different parts of the spectrum. The parts covered in the different views should be averaged, in order to reduce noise. The parts of the spectrum which are only present in only one of the views are simply kept. No noise reduction is performed in these spectral regions, since no complementary information is available. By introducing the assumption that low frequencies are common to all views, the estimated volume can be described as:

$$\tilde{V}(\mathbf{f}) = \begin{cases} \frac{1}{M}\sum_i V_i(\mathbf{f}), & \text{for } \mathbf{f} \leq \mathbf{f}_0 \\ max_i \left\{V_i(\mathbf{f})\right\}, & \text{for } \mathbf{f} \geq \mathbf{f}_0. \end{cases} \tag{4.12}$$

Figure 4.12 shows an example of result with this technique. In comparison to applying the same operator to all the spectrum, some of the details are lost at detriment of reducing speckle noise.

The main benefit of direct Fourier methods is its computational simplicity. However, as mentioned, the principal drawbacks of direct Fourier method are the global effects such as low frequency oscillation and ringing. Moreover, it is not trivial to discriminate signal from speckle in the spectral domain either, nor to guarantee the preservation of isolated features.



(a) (b) (c)

Figure 4.12: a) High Pass Maximum Frequency fusion image (32.66 dB), b) ROI, c) ROI on original image.

4.3.2 Wavelet Fusion

To overcome the aforementioned problems of direct Fourier methods, we are interested in a framework that offers a better spatio-spectral trade-off. Wavelet based fusion methods have been proposed to preserve spatial details of different acquisitions, and also, to remove speckle noise in ultrasound images. We pretend to combine these functionalities to both preserve details and reduce speckle, by exploiting the information from the different acquisitions [138]. Due to the compact support of the wavelets, strong reflectors are concentrated on few wavelet coefficients. On the contrary, noise is spread out among several coefficients, of smaller values [43]. Therefore, the goal is to enforce high-valued coefficients (features) and filter low-valued coefficients that are uncorrelated between views (noise).

In the multichannel radar imaging area, wavelet based fusion methods have been proposed to combine all spatial details of different acquisitions (e.g. [101]). Wavelet transforms of the images are computed, coefficients are fused following a certain rule into a single wavelet decomposition, and the fused image is obtained by taking the inverse wavelet transform of the fused coefficients. One possible fusion rule is to average coefficients of the approximation bands, and select the coefficients with the maximum absolute value for the detail bands. Other authors suggested an area-based maximum selection [33] and consistency checks to avoid spurious values [101].

Wavelet transforms have also been used to denoise images, particularly to despeckle ultrasound images (e.g. [61, 15]). A common technique consists in eliminating coefficients which have values under some threshold. Due to the compact support of the wavelets, spatially coherent signals concentrate on a limited number of coefficients, while noise spreads out. This is a commonly admitted hypothesis [87, 43, 61]. Coefficient elimination can be either via hard-thresholding (large values remain unchanged, values below the threshold become zero) or via soft-thresholding (coefficients are shrunk continuously, all values larger than the threshold are shifted by the threshold value). Weickert *et al.* [145, 113] demonstrated that the different thresholding techniques are actually equivalent to different diffusion-stopping functions of anisotropic diffusion. The threshold value may depend on the decomposition level and on the noise

and image statistics. An optimum threshold in the min-max sense for Gaussian additive white noise is proposed in [43], which is $t_l = \sigma_n\sqrt{2log(l)/l}$, where σ_n^2 is the noise variance and l the decomposition level. For speckle noise with Rayleigh distribution, it is proposed in [60] to use this threshold on the logarithm of the images, and then take the exponential of the fused result (homomorphic filtering). Soft-thresholding was applied among others in [61] for medical ultrasound images, adding a factor depending on image variance σ_x and an empirical factor k to the threshold: $t_l = k\sqrt{2\log(l)}\sigma_n^2/\sigma_x$. In general, σ_n^2 and σ_x^2 can be estimated from the image itself within a certain confidence range.

Wavelet basis

There is a long list of choices for the wavelet basis functions [105]. Longer wavelets tend to give a slightly better result in terms of speckle reduction, but they might oversmooth image details [64]. The number of decomposition levels needed to separate background texture from speckle also depends on the wavelet length, with more levels needed when a short wavelet function is used. In addition, the computational cost of the transform is nearly proportional to the length of the wavelet. We chose to use Daubechies wavelets, which are orthogonal, continuous, and have a good regularity. In particular we used length-4 Daubechies wavelet function and up to five decomposition levels ($L = 5$), which is a good compromise between support length and speckle robustness and proved to be efficient in our experiments. No major improvements were found with other classical wavelets. Other specific bases for ultrasound imaging could have been used, as proposed as in [15]. The comparison is left for future work.

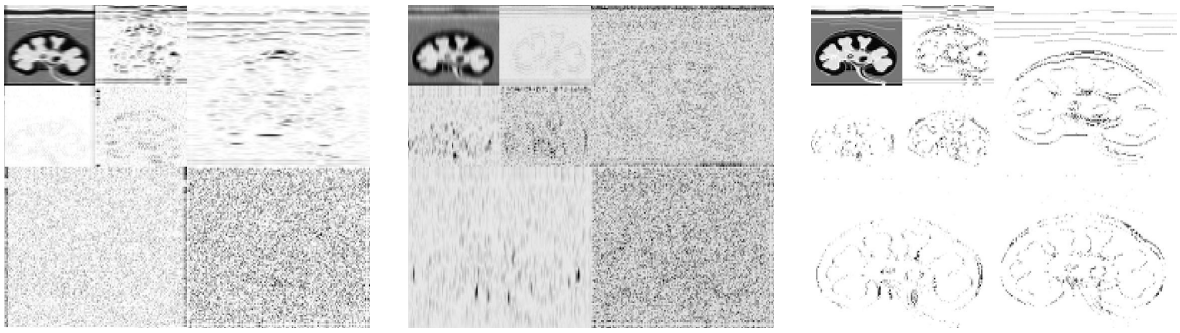
Figure 4.13 shows the wavelet decomposition of the example images, for length-4 Daubechies (up to two levels of decomposition, for clarity). For each image, the upper-left quadrant represents the low-frequency image and to the right, the horizontal decomposition; belw the vertical decomposition; and down-right the diagonal decomposition. It can be seen how the different acquisitions have the significant information in different quadrants and how the signal to noise ratio gets degraded at the high levels of decomposition.

Fusion rules

Our objective is to fuse the different views preserving spatial resolution and denoising, at the same time. We propose to use a combination of wavelet coefficients in order to keep the large coefficients and to eliminate the low-valued, uncorrelated coefficients. We define a fusion rule that fuses the coefficients along with a thresholding to reduce speckle.

We denote the wavelet transform of v_i by $w_i = \left\{w_i^1, \ldots, w_i^l, b_i\right\}$, containing the coefficients at each level and the low-pass component b_i. Our fusion approach consists in a simple averaging of the low-pass component b_i, and a maximum[1] of filtered wavelet

[1]Actually, in our experiments we performed a generalized average of the form $\tilde{a} = \left(\sum_{i=1}^{M} \frac{1}{M} a_i^{\beta}\right)^{1/\beta}$, with $\beta = 9$, which behaves close to the maximum operator. This was to avoid some minor artifacts



(a) (b) (c)

Figure 4.13: Length-4 Daubechies decomposition (2 levels) of: a) v_0, b) v_{90}, c) original image.

coefficients as follows:

$$\tilde{b}(\mathbf{y}) = \frac{1}{M}\sum_{i=1}^{M} b_i(\mathbf{y}) \tag{4.13a}$$

$$\tilde{w}^l(\mathbf{y}) = \max_i \left\{\alpha_i(\mathbf{y}) w_i^l(\mathbf{y})\right\} \tag{4.13b}$$

where $\mathbf{y}$ represents the wavelet transform coordinates. For $\alpha_i = 1$ (equal weights), this rule would preserve the largest components of each acquisition, thus preserving spatial details. In addition, in order to remove noise, we define the weight α_i as:

$$\alpha_i(\mathbf{y}) = \begin{cases} f\left(\frac{w_i^l(\mathbf{y})}{t_l}\right), & \text{for } \left|w_i^l(\mathbf{y})\right| < t_l \\ 1, & \text{otherwise.} \end{cases} \tag{4.14}$$

where f is a monotonous function such that $f(0) = 0$ and $f(1) = 1$. We used $f(x) = x^2$ in our tests. For Haar wavelet, such function is equivalent to a bi-weight Tukey function as diffusion stopping function [113], which has the edge-preserving properties described in Section 2.3.4. For Daubechies wavelet, such relationship has not been yet established. The value t_l is the threshold at level l defined in [61] as:

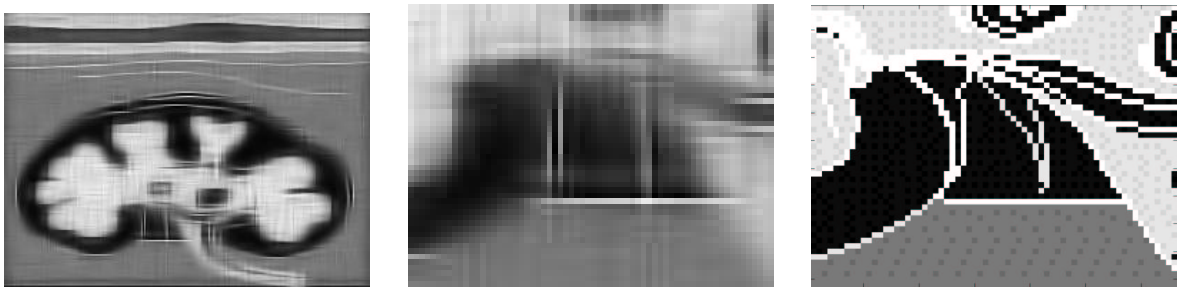
$$t_l = k\sqrt{2\log(l)}\,\sigma_n^2/\sigma_x \tag{4.15}$$

These weights will penalize in the fusion process the low valued wavelet coefficients, which are assumed to be noise. The parameter k controls the level of denoising, and has been set-up experimentally ($k = 3.5$), as a trade-off between denoising and detail preservation. The final image is computed by the inverse wavelet transform of $\tilde{w_i} = \left\{\tilde{w_i^1}, \ldots, \tilde{w_i^l}, \tilde{b_i}\right\}$.

Figure 4.14 shows the wavelet coefficient fusion results described in this section. Figure 4.15 shows the effect of the threshold level on the images. For $k = 0$ (Figure

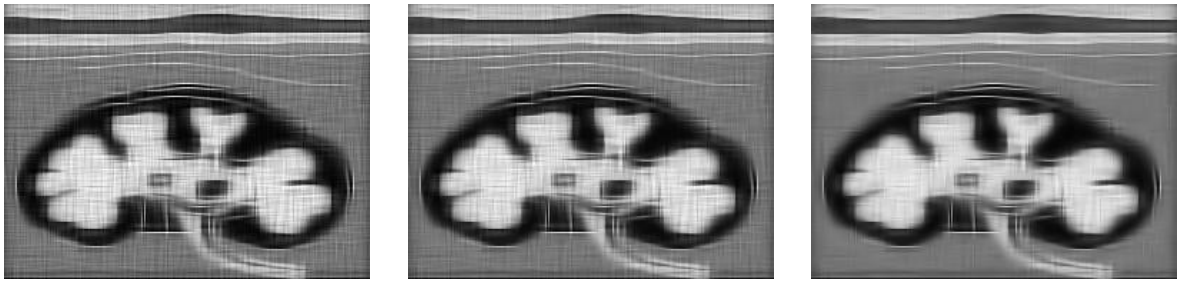
caused by abrupt transitions. For the sake of clarity, we express the operation as a maximum.

4.15 a), $t_l = 0$ and the image is not filtered. In this case, all the spectral content of the original images is kept, but all noise too. For higher values of k, wavelet coefficients corresponding to speckle are filtered out while spatial structures are still maintained if the threshold is kept within a certain range. Figure 4.16 shows the wavelet decomposition of the fused result for different threshold levels.



(a) (b) (c)

Figure 4.14: Wavelet fusion image (35.57 dB), b) ROI, c) ROI on original image.



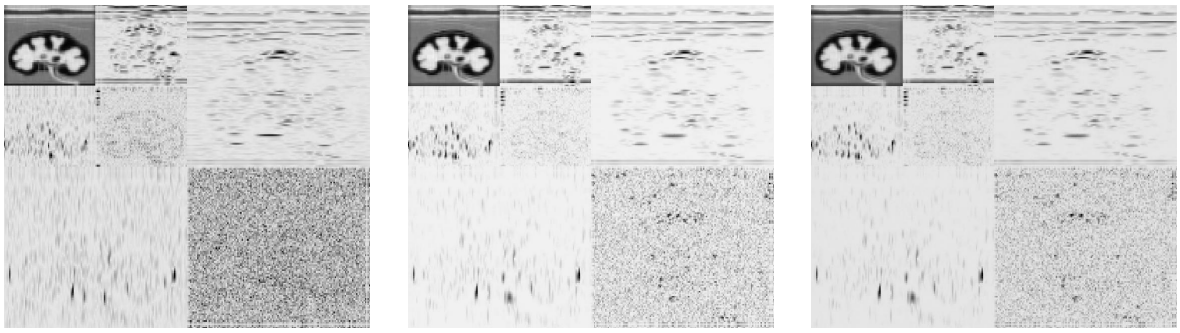
(a) (b) (c)

Figure 4.15: Wavelet coefficient fusion with different threshold levels. a) $k = 0$, no filtering (34.33 dB), b) $k = 2$ (35.41 dB), c) $k = 5$ (35.63 dB).

4.4 Comparison of the fusion methods

To evaluate the techniques presented in this chapter, we use the synthetic images presented in Section 3.4. Speckle noise has been added as $v_{noisy} = v + n \cdot \sqrt{v}$, where n is uniformly distributed random noise with zero mean. For this experiment, we have used $\sigma_n = \{0.005, 0.05, 0.5\}$, having previously normalized the images. The noisy image was blurred with an oriented Gaussian kernel with a variance $\sigma_k = \{2, 5, 8\}$ pixels at 0 and 90 degrees.

Figure 4.17 shows the fusion result of v_0 and v_{90} with the presented methods, namely: averaging, weighted averaging, generalized averaging, Fourier coefficient fusion, wavelet coefficient fusion and, as described in Chapters 2 and 3, multiview deconvolution. A zoomed view is shown in Figure 4.18.



(a) (b) (c)

Figure 4.16: Length-4 Daubechies decomposition for different threshold levels: a) $k = 0$, no filtering (34.33 dB), b) $k = 2$ (35.41 dB), c) $k = 5$ (35.63 dB).

Table 4.1 summarizes the quantitative results for this experiment. It contains the differences of the original image v with the reconstructed image $\tilde{v}$ in terms of PSNR (Section 3.4). *Noise*, in this context, is defined as the difference between the two images, accounting both for the differences due to spatial degradation and speckle noise.

σ_k	σ_n	v_0	v_{90}	avg	wa	ga	mf	wf	md
2	0.005	41.40	35.51	40.02	42.17	44.12	43.88	41.22	51.10
5		36.10	29.63	34.26	37.69	38.04	39.32	38.37	42.73
8		33.14	26.60	31.35	33.81	35.47	38.03	35.18	39.23
2	0.05	39.72	34.78	38.51	39.18	39.54	39.76	39.83	43.91
5		35.28	29.06	33.32	34.02	35.21	37.21	35.57	40.13
8		32.57	26.25	30.63	31.31	32.59	33.32	31.08	37.93
2	0.5	30.36	28.06	30.78	30.92	30.51	29.98	30.67	31.17
5		29.56	25.87	28.53	28.17	28.16	28.84	29.55	30.13
8		28.20	24.02	26.84	25.21	25.18	25.71	29.03	29.53
average computing time (ms)				10	251	401	376	471	1272

Table 4.1: Peak Signal to Noise Ratio (dB) for simulation data, for degradations σ_k and speckle noise standard deviation σ_n, and average computing time (ms). Methods: average (avg), weighted average (wa), generalized average (ga), maximum frequency (mf), wavelet fusion (wf) and multiview deconvolution (md).

The most important observation is that multiview deconvolution is the only outstanding method. It shows better delineation, preserves as many features as others and shows a relatively good noise reduction. Indeed, it quantitatively performs best in all conditions of noise power and spatial degradation. The differences are greater in best case conditions (low noise $\sigma_n = 0.005$, small blur $\sigma_k = 2$), but also in worst case conditions (high noise $\sigma_n = 0.5$, large blur $\sigma_k = 8$). Note that considering the PSF of the system automatically collects the features visible at each view. In comparison

to other compounding strategies where the angle of incidence is considered, for this technique it is the consideration of the PSF which makes the difference between the contributions of each of the acquisitions.

For the methods described in this chapter, it is important to mention the denoising capabilities of wavelet methods. Actually, it is the only one which explicitly filters out noise, and the results can be appreciated at the bottom of Figure 4.18.e. While the denoising strategy of the other proposed methods consists in averaging the different uncorrelated views, the wavelet framework enables to better isolate edges from homogeneous areas, facilitating its elimination. Indeed, in high noise conditions its performance is close to the regularized multiview deconvolution technique. In all cases, it performs better than the spatial domain techniques. On the other hand, small artifacts appear, especially close to edges in the reconstructed image.

It is also interesting to mention the performance of the generalized average technique, which shows a good spatial resolution preservation. Figure 4.18.c shows how it slightly improves the definition of details while still keeping most of features.

Fourier coefficient fusion shows an overall performance comparable to other fusion techniques, but shows a higher level of high-frequency noise. Although visually perceptible, it is not reflected into the overall metrics.

Weighted averaging improves local contrast when compared to averaging, as perceptible in Figure 4.18.b. However, spatial resolution and denoising are similar to the simple averaging method.

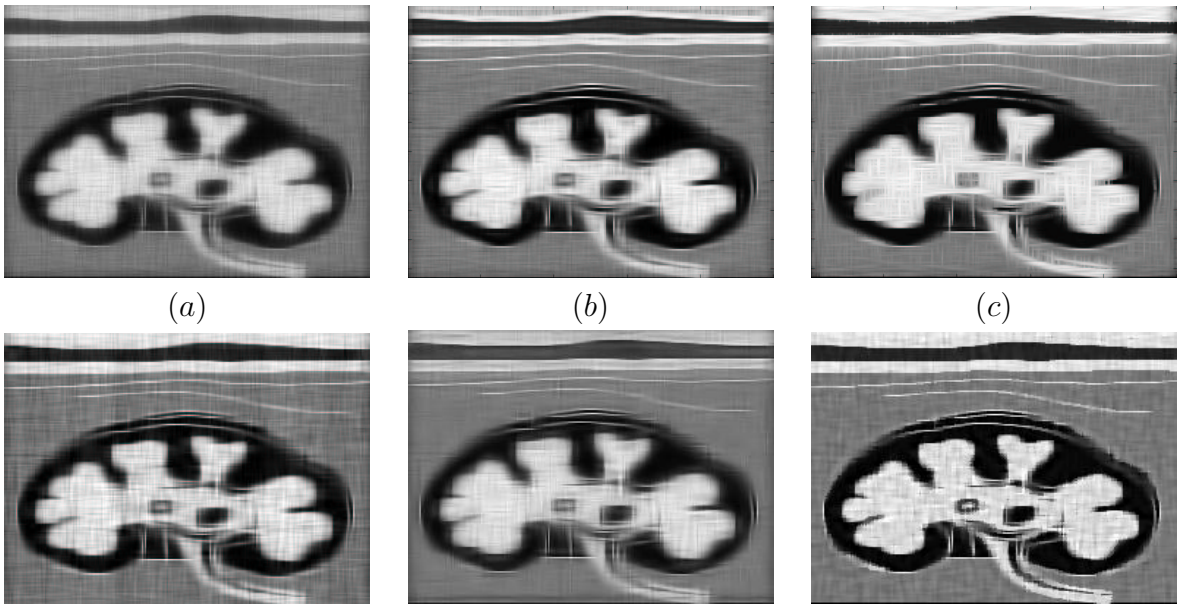


Figure 4.17: Fusion of synthetic images of Figure 3.1 a and b. a) average (33.32 dB), b) weighted average (34.02 dB), c) generalized average (35.21 dB), d) Fourier fusion (37.31 dB), e) wavelet fusion (35.57 dB), f) multiview deconvolution (40.13 dB).

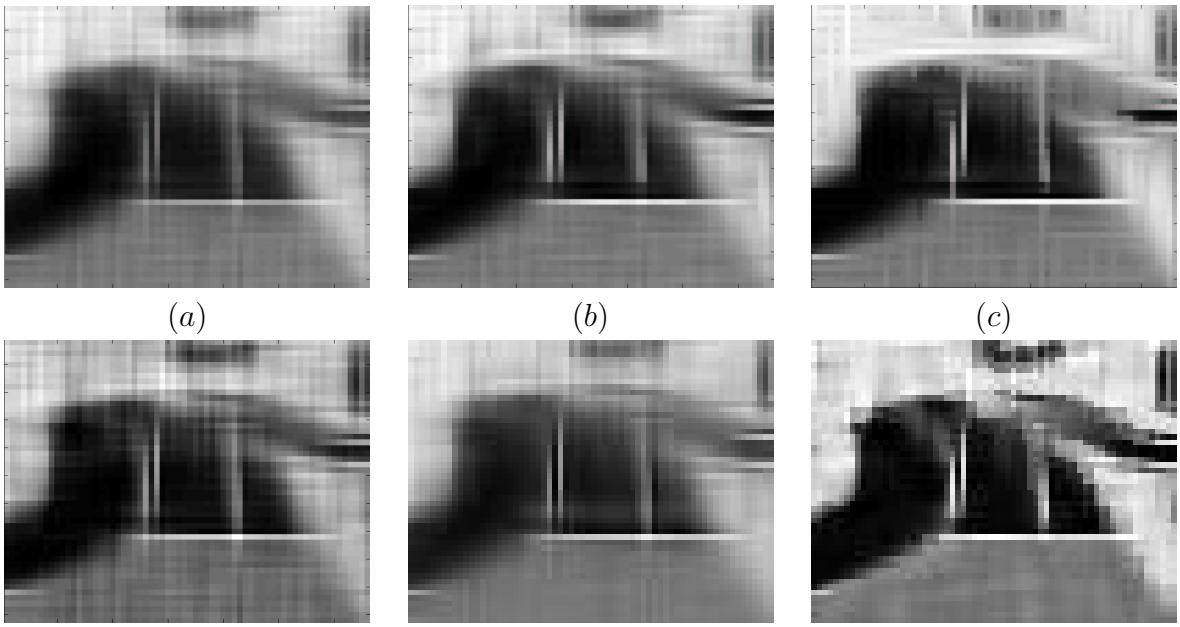


Figure 4.18: Zoom on fusion of synthetic images of Figure 3.1 a and b. a) average (33.32 dB), b) weighted average (34.02 dB), c) generalized average (35.21 dB), d) Fourier fusion (37.31 dB), e) wavelet fusion (35.57 dB), f) multiview deconvolution (40.13 dB).

Computational times for this example image (8 bits, 186x250 pixels) are also presented in Table 4.1. The average method is the fastest one (10 ms), followed by weighted average (215 ms), wavelet fusion (471 ms) and multichannel deconvolution (1272 ms). For the later, this time includes both the estimation of the system's PSF and the volume reconstruction. This computational time figures should be regarded as relative since they depend on many parameters such as the platform and the coding language. Therefore it is interesting to notice that taking averaging as reference (x1), the cost of the techniques is as follows: weighted average (x22), wavelet fusion (x47) and multichannel deconvolution (x127).

4.5 Conclusions

In this chapter, we have reviewed some methods to fuse different views of an image, from averaging to wavelet coefficient fusion. Those were based on a fusion approach, in the sense that the reconstructed volume is made of pieces of information from each view. We have introduced several refinements in order to achieve our imaging goals. However, these techniques, on synthetic images, underperformed the regularized multiview deconvolution method. The reason is that this method actually reverts the degradation process, compensating the effect of the PSF of the acquisition process, while the regularization step avoids noise amplification.

The methods presented in this chapter do not equalize the degradation process. On the other hand, they are faster since they do not need to estimate the PSF, they have a closed-form solution and in general less parameters to estimate. Indeed, they preserve features present only in some of the views, which may be sufficient for some applications. Spatial domain techniques show a relative low computational cost, but they suffer from the difficulty to discriminate features form speckle. The list of methods is not exhaustive. Notably, fusion schemes based on pyramidal decomposition [34], such as Oriented Laplacian decomposition, are of great interest for this application. Our results with these techniques show a performance close to those presented here, thus they have not been included in this study.

Wavelet framework, as proved in the literature, shows to be an efficient denoising framework. Indeed, it could be combined with the multiview deconvolution technique. Many of the functionalities described in this chapter (measures of saliency, discrepancy and speckle) could be integrated into the multiview framework to improve its performance.

With this chapter, we finished to describe the methods we developed to combine the different acquisitions. Next chapters are dedicated to apply this methodologies to real data with clinical applications.

CHAPTER 5

3D Breast Ultrasound Imaging

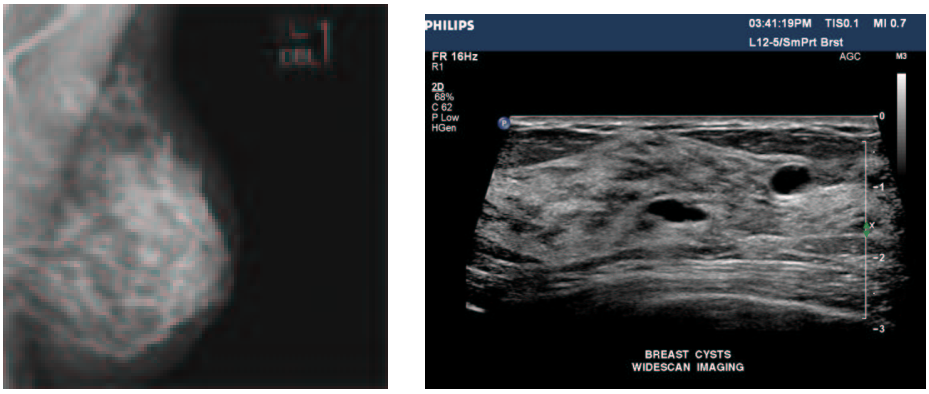
5.1 Introduction

One woman in eight will develop breast cancer in her life time [4]. In year 2000, over 1 million women were diagnosed with new cases of breast cancer worldwide and 373,000 died from this disease [50]. However, if detected early, the five-year survival rate exceeds 95%. Therefore, the importance of breast imaging systems devices capable of diagnosing this disease is vital. Indeed, the development of medical imaging systems has contributed to the reduction in breast cancer mortality rates in recent years [4].

X-rays mammography is among the best early detection methods and is currently used as a screening tool in clinical routine. It permits the detection of small structures, such as microcalifications, and can image most types of breast. However, it is invasive due to the X-ray radiation and cannot diagnose all breast cancer pathologies. This is specially important for breast masses, since it has a low contrast resolution which does not permit to image fluid cysts [111]. Ultrasound echography, on the contrary, has an excellent contrast resolution and is non-invasive. Indeed, devices are portable and exams are relatively cheap. On the other hande, it has a limited spatial resolution [69], it cannot detect microcalifications, it cannot image dense fat breasts and the image quality is dependent on the practitioner's skills [111]. Figure 5.1 shows an example of standard X-ray mammography and ultrasound exams.

While current ultrasound clinical practice is performed with 2D imaging systems, 3D imaging systems can improve substantially the diagnosis abilities [96]. Masses and tissues can be completely captured and shape parameters accurately quantified. Indeed, if the whole breast is scanned, the exam is less dependent on the practitioner, since all features are captured within examination time. However, up to date, there are no planar array 3D acquisition systems in the market suited for imaging the whole breast. We use a linear probe embedded in a robotic platform which scans the whole breast tissue. Other approaches have been proposed, such as hand-held linear array attached to a linear slider [96], or free-hand systems with a position sensor [164][129].

Obtaining 3D acquisitions of the breast is a challenging task. Due to the limited elevational resolution of linear arrays (see Appendix A) the obtained volumes have a highly anisotropic resolution, which limits the spatial resolution of the system. In this



(a) (b)

Figure 5.1: Example of breast tissue exams with a) x-ray mammography and b) ultrasound echography, from [6]

chapter we propose a technique to improve the performances of the system [140]. By scanning the breast tissue in different directions, views which have a different point spread functions can be obtained. By combining them, the *spatial resolution* will be improved. Moreover, these different views have uncorrelated speckle patterning. By combining them, the SNR will be increased, improving the *contrast resolution*. In other words, the acquisition of these different views permits us to perform 3D deblurring and 3D spatial compounding. This application is therefore entirely relevant with respect to the methodological developments described in the previous chapters.

We give a brief introduction to breast cancer in Section 5.2. The acquisition procedure is explained in Section 5.3. In Section 5.4 we will study the characteristics of the data, in order to validate the models used for noise, signal and the point spread function. Non-rigid registration techniques which ensure proper alignment of the different views are described in Section 5.5. The multiview techniques developed in Chapters 2 and 3 are applied to obtain high resolution 3D scans of the breast tissue. Results on *in vivo* data are reported in Section 5.6, and conclusions drawn in Section 5.7.

5.2 Clinical needs

5.2.1 Breast cancer

The term "breast cancer" denotes a wide, heterogeneous range of malignant diseases of breast tissues. Adenocarcinoma is a general type of cancer that starts in glandular tissues anywhere in the body. Nearly all breast cancers start in glandular tissue of the breast and are, therefore, adenocarcinomas. The two main types of breast adenocarcinomas are ductal carcinomas and lobular carcinomas. There are also several cellular subtypes, some of which have important implications for prognosis and treatment, and different types can coexist simultaneously. The most common breast malignancies are summarized in Table 5.1 [137]. Figure 5.2 shows a diagram of normal breast structures.

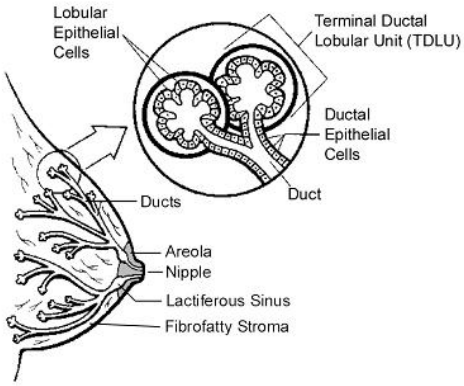


Figure 5.2: Diagram of normal breast structures

Classification	**Type**	**Sub-type**
Non-invasive Epithelial Malignancies	Ductal Carcinoma In Situ (DCIS)	Comedo Solid Cribriform Papillary Micropapillary
	Lobular Carcinoma In Situ (LCIS)	
Invasive Epithelial Malignancies	Invasive Ductal Carcinoma(IDC)	Scirrhous Transitional Cellular IDC + DCIS
	Invasive Lobular Carcinoma (ILC)	
	Medullary Carcinoma	
	Mucinous (Colloid) Carcinoma	
	Papillary Carcinoma	
	Tubular Carcinoma	
	Paget's Disease	
	Inflammatory Carcinoma	
Non-epithelial Malignancies	Sarcoma	Cytosarcoma Phyllodes Fibrosarcoma Liposarcoma Angiosarcoma
	Lymphoma	

Table 5.1: Most common malign breast pathologies [137]

Most breast diseases are benign, and most breast symptoms are caused by fibrocystic changes. Benign breast tumors such as fibroadenomas or papillomas are abnormal growths, but are not malignant and cannot spread outside of the breast to other organs. The most common benign breast conditions are summarized in Table 5.2 [137].

Type	**Sub-type**
Fibrocystic Change(FCC)	Cysts Duct ectasia Ductal hyperplasia Lobular hyperplasia Fibrous mastopathy Microcystic mastopathy Focal fibrosis Adenosis
Fibroadenoma	Fibroadenoma Proliferative Phyllodes
Mastitis	Puerperal Non-puerperal
Other	Papilloma Lipoma Hamartoma Fibromatosis

Table 5.2: Most common benign breast pathologies [137]

There exist some classification methods to assess the malignancy of breast disease. Among others, BI-RADS^{tm} (Breast Imaging Reporting and Data System) [1] is one of the most popular ones. Many parameters are taken into consideration to evaluate the disease malignancy, such as size, contour shape or mass density. According to their values and relationships, BI-RADS^{tm} classifies the lesion into five categories, from negative finding to high confidence on malignancy (Table 5.3). Typical indications of malignancy are spiculated and elongated masses, e.g. an invasive scirrhous ductal carcinoma.

Category 0:	Need Additional Imaging Evaluation
Category 1:	Negative
Category 2:	Benign Finding
Category 3:	Probably Benign Finding; Short Interval Follow-Up Suggested
Category 4:	Suspicious Abnormality; Biopsy Should Be Considered
Category 5:	Highly Suggestive of Malignancy; Appropriate Action Should Be Taken

Table 5.3: BI-RADS^{tm} Classification.

5.2.2 Imaging needs

The imaging system has to reveal the parameters that permit differentiating benign and malignant diseases, following the BI-RADS classification or other.

Contrast resolution

For the breast imaging application, it is important to have a good contrast resolution. Contrast resolution refers to the ability to distinguish a mass from its borders, or the mean density of two masses. If the image is corrupted by noise, the density of the mass may not be distinguished from its neighbors. Figure 5.3 illustrates this concept with a synthetic example of two masses which have a similar density value, but are hardly distinguishable under noise condition. The objective of our system is to improve the SNR in order to improve the contrast resolution, to better appreciate masses.

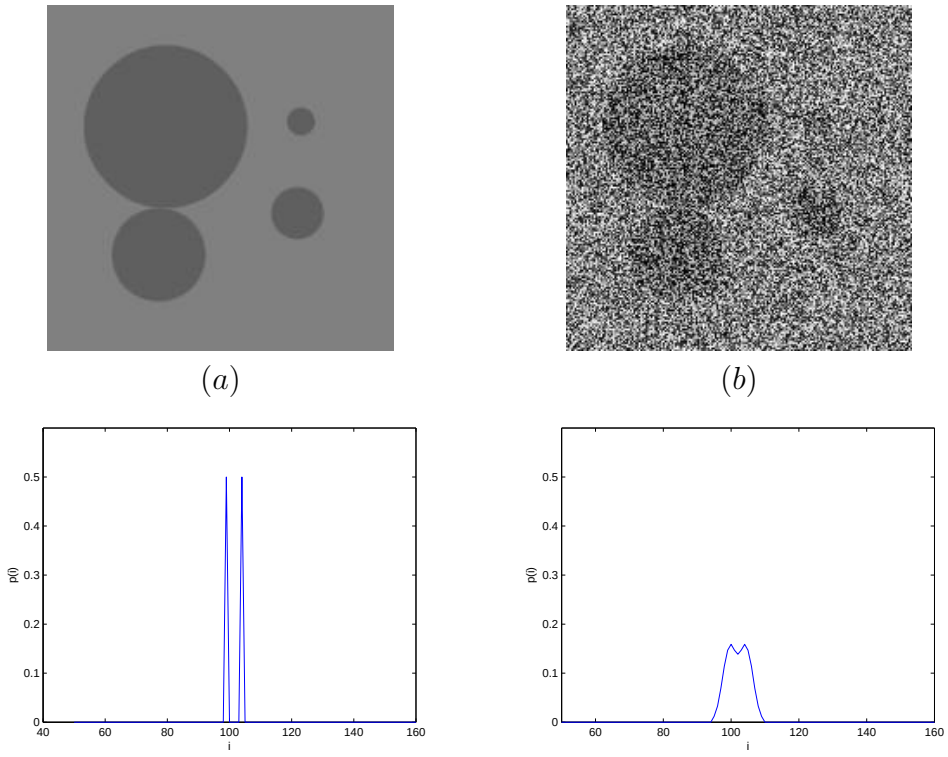


Figure 5.3: Contrast resolution is greatly reduced by noise: a) original simulated masses, b) image with additive Gaussian noise, c) histogram of (a), d) histogram of (b).

Spatial Resolution

Spatial resolution refers to the ability of the system to distinguish closely spaced reflectors. In order to appreciate mass borders and small fibers, a good spatial resolution

is needed. Otherwise, there is a loss in details which can hide elements of diagnosis. Figure 5.4 shows an example of a spiculated mass and a round-shaped mass, which are hard to distinguish under blurry conditions.

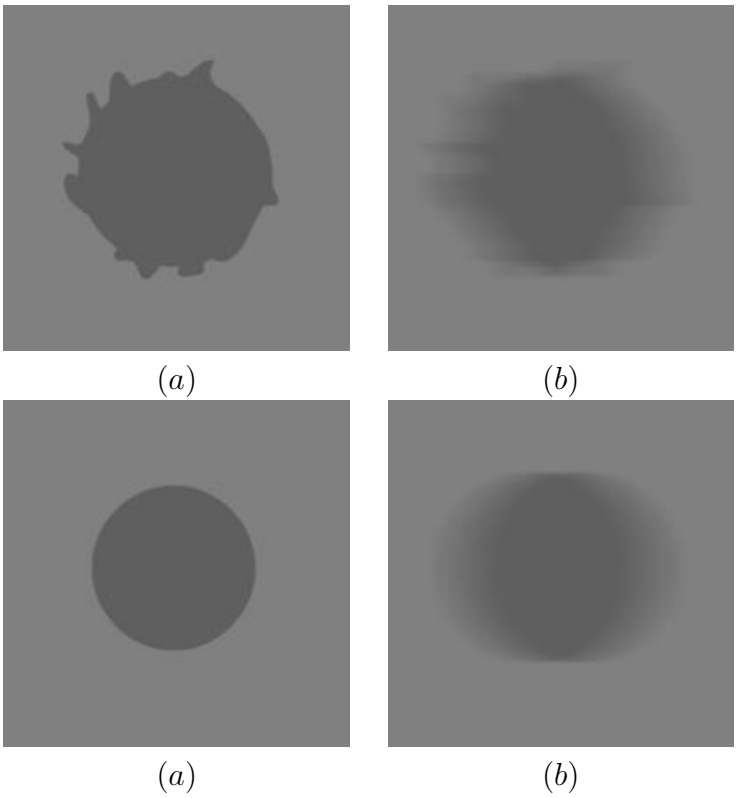


Figure 5.4: Spatial resolution is reduced by the effect of the point spread function: a) spiculated simulated mass of 5mm of diameter, b) spiculated mass blurred with a Gaussian PSF (σ = 1mm), c) round-shaped simulated mass of 5mm of diameter, d) round-shaped simulated mass blurred with Gaussian PSF (σ = 1mm).

In order to detect microcalcifications, a spatial resolution of about $50\mu m$ is needed. This is far below the resolution of current conventional ultrasound systems, which show a typical spatial resolution of about 300x1000x300μm in lateral, elevational and axial dimensions. However, even if the smallest microcalifications are not visible, by improving the spatial resolution the microcalifications will be detectable at an earlier stage of their evolution.

Practitioner Independence

Current systems depend on the skills of the practitioner. Masses can be missed or not acquired in its larger or right view in 2D images, which is not the case in 3D. By using a mechanical platform, the dependency on the practitioner's skill is greatly reduced. Exams are made in a systematic way which avoids missing interesting features. Reducing variability lead to more robust performance on computer aided diagnosis.

5.3 Data Acquisition

5.3.1 Methods

In order to obtain 3D acquisitions, a linear array was embedded in a two-dimensional robotic platform, as schematized in Figure 5.5. The linear array layed in th xz plane, insonifing in the y direction. The tissue was scanned by linearly translating the probe along the plataform. Regular 2D ultrasound images were obtained at each position of the probe, constituting a series of slices of the volume of interest. The platform has three degrees of freedom (lateral translation x, elevation translation z and probe rotation θ), which permit scanning the tissue in any direction. Elevational tilt of the probe was set to 90 degrees, obtaining images exactly perpendicular to the platform (xz plane). The breast tissue was compressed onto the scanning platform and the patient was asked not to move during all scans.

In order to obtain different views, we scanned the breast tissue along different directions, in increments of 45 degrees, as depicted in Figure 5.6. Each linear scan takes approximately 15 seconds, resulting in a total scanning time of 1 minute. These different views have uncorrelated speckle patterning and rotated point spread functions, permitting to improve the SNR and overcome the limited elevational resolution by combining them. We will denote these views as v_0, v_{45}, v_{90} and v_{135}, as short notation for scalar-valued function $v(\mathbf{x}) : \Omega \rightarrow \mathbb{R}, (\Omega \subset \mathbb{R}^3)$ where $\mathbf{x} = (x, y, z)^T$ represents the position in the space, with x referring to the lateral dimension, y to the axial dimension and z to the elevational dimension.

The B-scan refers to the conventional two-dimensional ultrasound image of (xy plane for v_0), the elevation scan (E-scan) refers to the perpendicular plane to the B-scan and parallel to the scanning direction (yz plane for v_0), and the C-scan refers to the plane perpendicular both to the B-scan and elevational planes (xz plane for all acquisitions), as shown in Figure 5.7. Figure 5.8 shows an example on a fruit phantom (Section B.3) of a B-scan, E-scan and C-scan for v_0 and v_{90}. Notice the different point spread functions of these planes resulting in a larger blurring effect in the E-scan and C-scan.

The robotic platform is currently installed at the University of Chicago Hospital, IL, USA for clinical validation studies.

5.3.2 Ultrasound Acquisition System

The ultrasonic probe was a Philips L12-5 50 mm broadband linear array with 192 elements and operating frequency range from 5 to 12 Mhz, on a Philips HDI5000 system. Two-dimensional images were obtained each 0.2 mm, triggered by the position encoder. Each 2D slice was acquired with spatial compounding (SonoCT^{tm}) with three angulations of -8.5, 0, and 8.5 degrees. The width of the region scanned was 5 cm, at a spacing between acquisitions of 0.2 mm, the depth depended of the imaged target, varying from 2.5 cm to 5 cm, and resolution varying from 0.05 to 0.1 mm. A final volume was linearly interpolated onto a grid of voxel size 0.2x0.1x0.2 mm, for lateral, axial and elevational dimensions respectively.

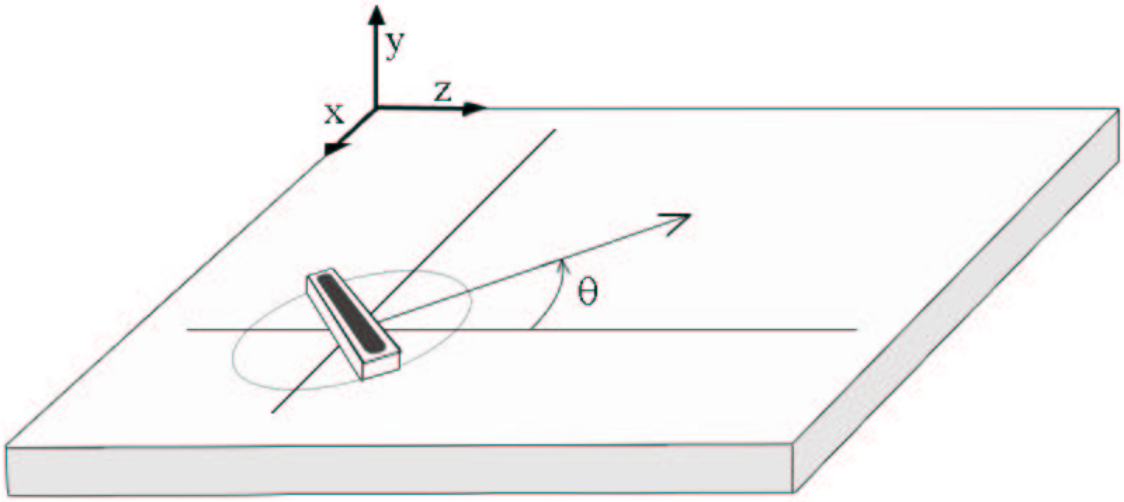


Figure 5.5: Robotic platform for 3D ultrasound breast scanning, with three degrees of freedom: lateral axis (x), elevational axis (z) and probe angulation (θ). Dimension y represents the axial direction, the depth of the ultrasonic scan.

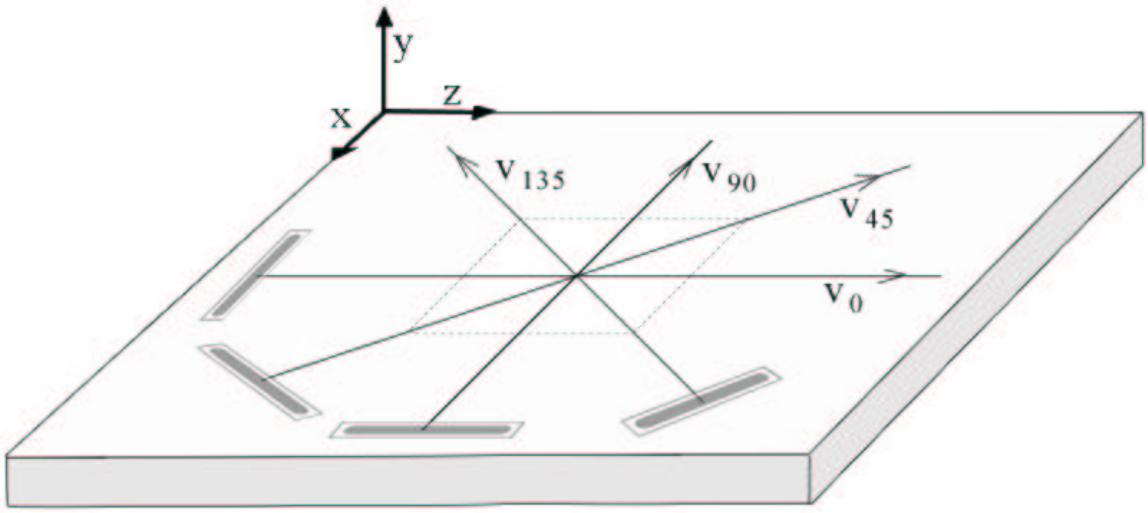


Figure 5.6: Multiview scanning strategy. The tissue volume within the dashed line is scanned in four directions, at 0, 45, 90 and 135 degrees, to obtain views with complementary point spread functions and uncorrelated speckle paterning.

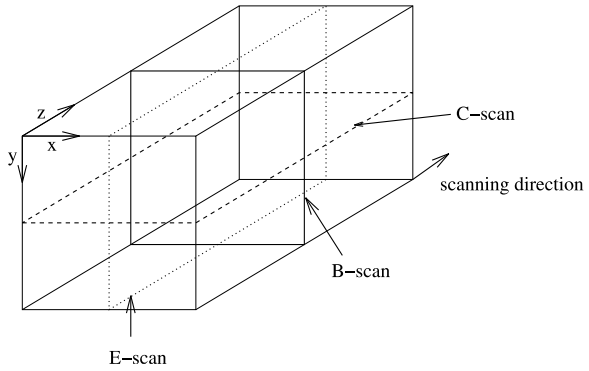


Figure 5.7: Definition of B, E and C scan for 3D breast acquisition for v_0.

5.4 Data Characterization

In this Section we characterize the statistics of the signal and noise of the acquired data, in order to validate the signal and noise models. These statistical properties will be used as parameters in the fusion algorithms, as indicated in Section 5.6.1.

5.4.1 Point Spread Function

The point spread function of the system depends on the probe characteristics (Section A.3.1). Briefly, the axial resolution depends on the ultrasound pulse length, and the lateral and elevational resolutions depend in the aperture of x and z dimensions respectively. The phantom "bubbles" (Appendix A) was used to characterize the point spread function of the system. It is made of small isotropic small bubbles that are close to be punctual, thus they can be considered as individual reflectors. Figure 5.9 shows the B-scan and C-scan for v_0 and E-scan and C-scan for v_{90}. Figure 5.10 shows a zoom on the C-scan for one of these bubbles for all four acquisitions, which highlights the anisotropy of the acquisitions due to the limited elevational resolution.

Figure 5.11 displays the profile in the lateral, axial and elevational directions for a bubble in v_0, which illustrates the worse resolution in the elevational plane, besides the apparition of secondary lobes due to the nearly squared elevational aperture. In the axial profile, an artifact due to the construction of the phantom can be observed. There is a second local maximum after the main PSF, due to the clutter caused by the bubble nature of the impulsional particle.

In order to quantify the size of the point spread function from the phantom, the full width at half maximum (FWHM) [157] was measured for all bubbles as a function of x, y and z. The FWHM is the distance between the points at which the PSF reaches half its maximum value, defined as:

$$FWHM = |\mathbf{x}_1 - \mathbf{x}_2|\,, h(\mathbf{x}_1) = h(\mathbf{x}_2) = \max(h(\mathbf{x}))/2 \tag{5.1}$$

The goodness-of-fit of the Gaussian model and the measured data is measured by

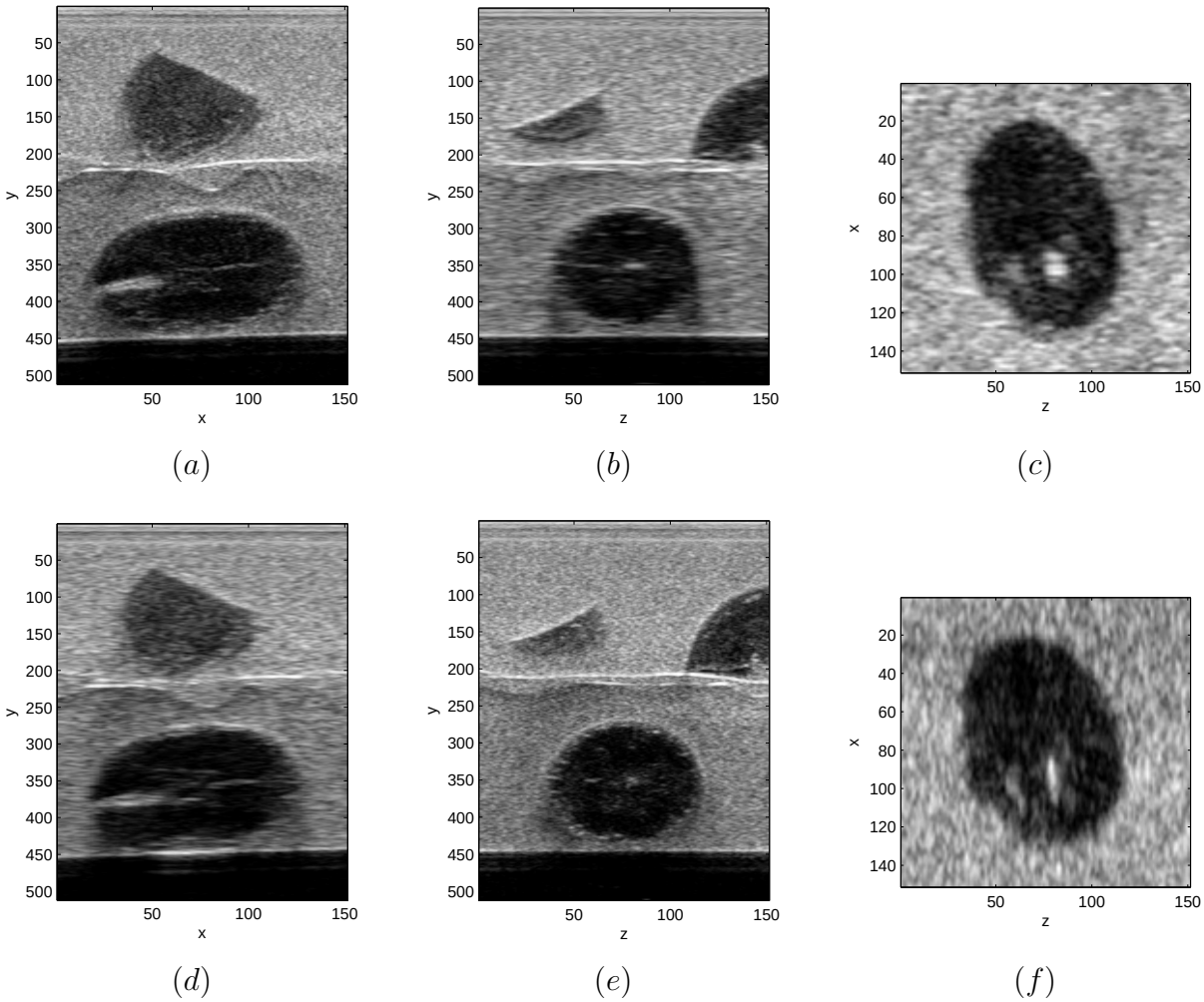
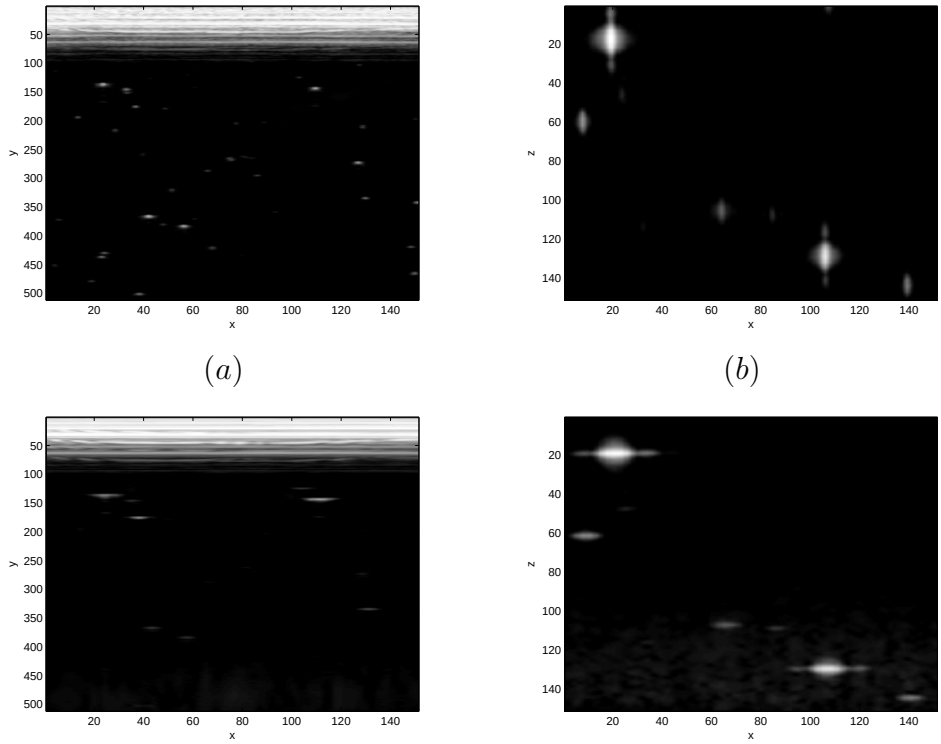


Figure 5.8: Slices of fruit phantom 3D acquisition, zoom on a 3x5x3 cm area, voxel size 0.2x0.1x0.1 mm. For v_0 (upper row): a) B-scan ($z = 60$), b) E-scan ($x = 100$), c) C-scan ($y = 350$). For v_{90} (lower row): d) E-scan ($z = 60$), e) B-scan ($x = 100$), f) C-scan ($y = 350$). Due to the point spread function of the system, the best spatial resolution is achieved in the B-scan, which corresponds to the xy plane for v_0 and to the zy plane for v_{90}. The objective of the mutliview restoration algorithm is to combine these different views to obtain a high-resolution volume.

the ρ^2 metric, defined as:

$$\rho^2(A, B) = \frac{\text{cov}(A, B)}{\sigma(A)\sigma(B)} \tag{5.2}$$

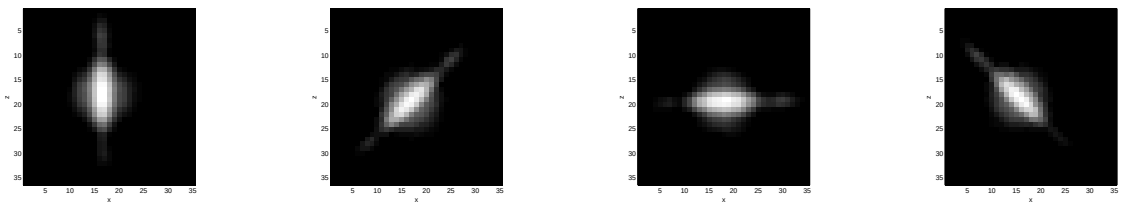
where $\text{cov}(A, B)$ is the covariance between the two data sets, and $\sigma(\bullet)$ is the standard deviation of a particular data set. This metric takes the value 1 for a perfect fit and decreases toward zero as the fit gets worse. Although there might be more appropriate fitting tests for this kind of data (e.g. Kolmogorov-Smirnov [36]), we chose this method because it gives a normalized result.



(a) (b)

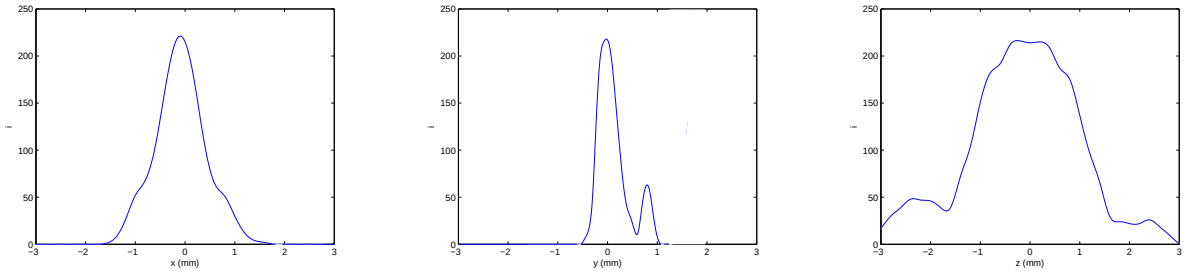
(c) (d)

Figure 5.9: Bubble phantom to measure the point spread function: a) B-scan($x = 80$) of v_0, b) C-scan($y = 400$) of v_0, c) E-scan($x = 80$) of v_{90}, d) C-scan($y = 400$) of v_{90}.



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

Figure 5.10: C-scan slice of the point spread function, for a) v_0, b) v_{45}, c) v_{90} and d) v_{135} (pixel size 0.2x0.2mm).



(a) (b) (c)

Figure 5.11: Point spread function profile of v_0 for bubble centered at (x=10.40, y=26.39, z=5.40 mm). a) Lateral, b) axial, c) elevational profiles. Notice the secondary peak in the axial profile due to clutter, and the secondary lobes in the elevational profile.

Table 5.4 reports the results for this metric. As expected, the PSF is largest in the elevational plane (2.34 mm), followed by the lateral plane (1.10 mm) and the axial plane (0.61 mm). Table 5.4 also shows the fit (ρ^2) of the PSF to a Gaussian model, showing a good fit overall. The lateral profile shows the best fit to a Gaussian model (0.9884), followed by the elevational profile (0.9606) and the axial profile (0.9441). This mismatch is due to the presence of clutter and the secondary lobes. If these parts were not considered, the goodness-of-fit would improve to 0.9711 and 0.9763 for the elevational and axial dimensions respectively. Also, there is a slight discrepancy between the values obtained by the Gaussian fit and the FWHM metric, which should be related as $FWHM = 2\sqrt{2\ln 2}\sigma = 2.3548\sigma$. Due to the clutter artifact and the secondary lobes, the Gaussian fit tends to overestimate the size of the PSF.

	FWHM (mm)	σ	$2\sqrt{2\ln 2}\sigma$	ρ^2
lateral	1.10±0.29	0.55±0.16	1.30	0.9884
axial	0.61±0.12	0.28±0.51	0.66	0.9441
elevational	2.34±0.60	1.32±0.31	3.11	0.9606

Table 5.4: Point spread function size (FWHM in mm), for lateral, axial and elevational directions on bubble phantom, and fitted Gaussian model, σ (mm), and goodness-to-fit (ρ^2).

Figure 5.12 shows the dependency of the size of the point spread function with the depth (y axis), especially for the elevational plane. This is due to the fixed focalization of the acoustic lens of the probe. The dependency on depth of the PSF in the lateral and axial planes is smaller. The goodness of fit of the Gaussian model was uniform for all depths.

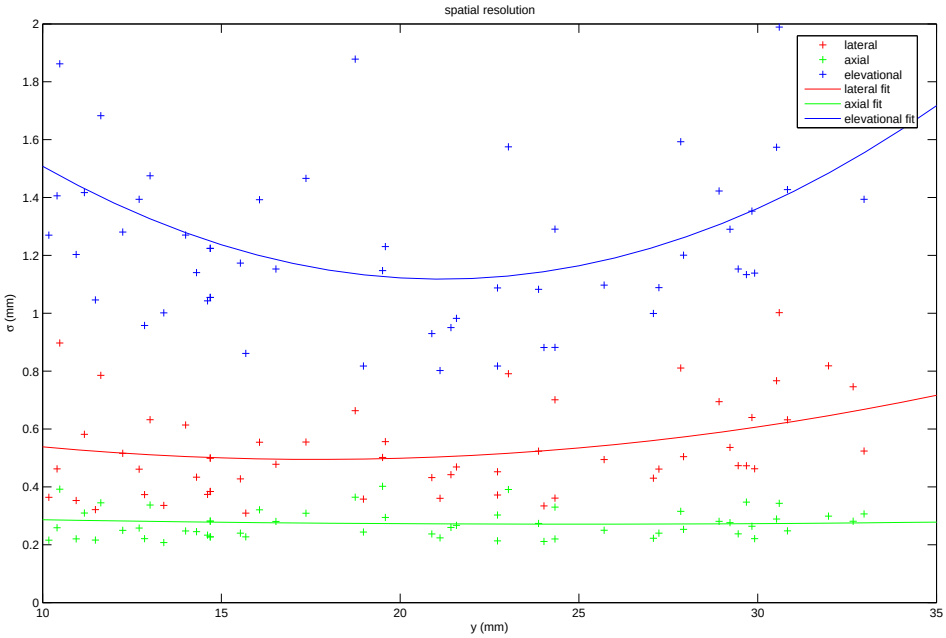


Figure 5.12: Point spread function size (σ) as a function of depth, for lateral (red), axial (green) and elevational (blue) directions. Points represent individual bubble sizes, solid line represent quadratic fit.

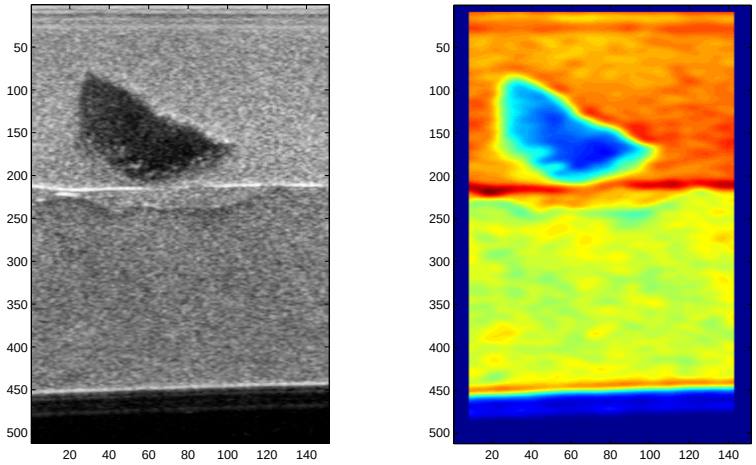
5.4.2 Speckle Statistics

In this section we examine the hypothesis that speckle in log-compressed ultrasound images follows a colored Gaussian additive noise model [79].

Intensity Distribution

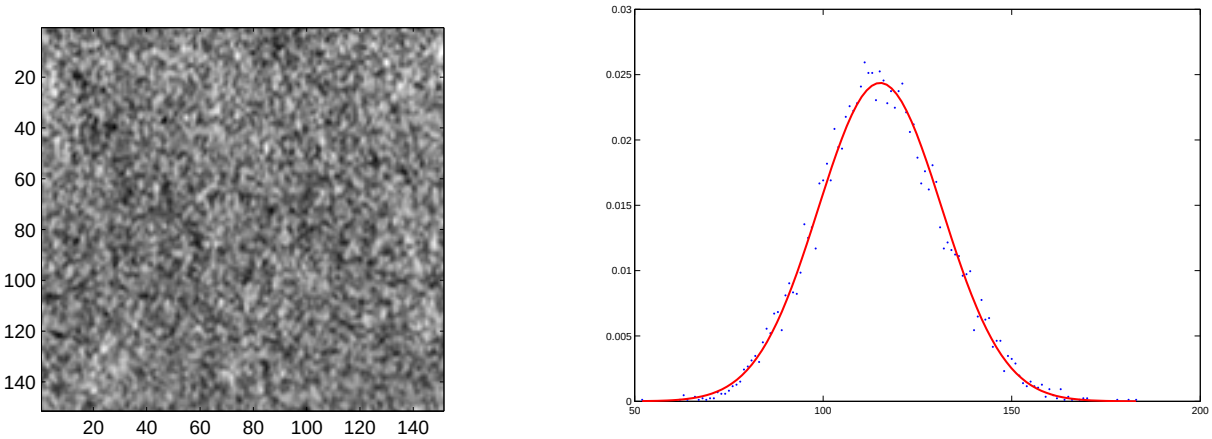
We examined the intensity distribution of speckle in order to understand its statistical behavior and verify our model. Uniform areas, where supposedly variations are caused only by speckle patterning, were selected as areas where gradient magnitude is the lowest. Figure 5.13 shows the three main areas used in the fruit phantom, where three regions can be clearly differentiated. Figure 5.14 shows a uniform area of speckle along with its histogram and its Gaussian fit - in a maximum-likelihood sense.

Table 5.5 shows the fitted parameters for B-scan planes, C-scan planes, E-scan planes and the 3D volume. Results showed an extremely good fit with Gaussian distribution, $\rho^2 > 0.99$ for $\mu = 110$ and $\mu = 150$. The average fit for $\mu = 35$ is lower, $\rho^2 = 0.89$, because as the mean average value is closer to zero, its distribution gets closer to a Rayleigh distribution. The variance of the noise power for this phantom is larger for lower values since their area is smaller than others. The variance is larger in all cases for the C-scan plane, due to the blurring effect of the point spread function. From this data, we can make a rough estimation for at which point the noise



(a) (b)

Figure 5.13: a) B-scan plane of fruit phantom, b) colored local mean of a), where three main areas can be differentiated as red ($\mu = 145$), green ($\mu = 110$) and blue ($\mu = 35$).



(a) (b)

Figure 5.14: a) Uniform area corrupted with speckle ($\mu = 110$) and b) its histogram (points, blue) and Gaussian fit (solid, red, ρ^2=0.9942)

distribution is no longer Gaussian. Considering the variance of the noise, the lower tail of the distribution will be higher than zero for values of $\mu \geq 50$, which represents a 80% of the gray scale range. For lower values, the fit is worse. This result is consistent to assume a Rician model for the speckle intensity distribution (Appendix A) - for low levels of signal, the distribution is Rayleigh, and for higher, is close to Gaussian.

	$\bar{\sigma_B}$	$\bar{\sigma_E}$	$\bar{\sigma_C}$	σ_{3D}	ρ^2_B	ρ^2_E	ρ^2_C	ρ^2_{3D}
$\mu = 35$	3.43±1.5	3.32±1.5	3.21±1.9	3.31±1.7	0.918	0.850	0.875	0.891
$\mu = 110$	3.99±0.6	3.87±0.7	3.90±1.3	3.92±0.9	0.994	0.995	0.985	0.992
$\mu = 145$	4.22±0.4	4.17±0.3	3.76±1.6	4.13±1.2	0.997	0.997	0.995	0.995
average	3.88	3.78	3.65	3.76	-	-	-	-

Table 5.5: Speckle intensity standard deviation for v_0 on the fruit phantom, along B, E and C planes, and fit to Gaussian distribution.

Relationship between Mean and Standard deviation

The relationship between mean and standard deviation was evaluated for all mean gray-values in uniform areas on phantom and *in vivo* data. The standard deviation was found almost constant for all mean values, with a slight linear dependency. In other words, the noise variance is constant throughout all the range of values of log-compressed ultrasonic B-scan, E-scan, C-scan and 3D ultrasound data as acquired in our application. This justifies the use of an additive noise model rather than a multiplicative one.

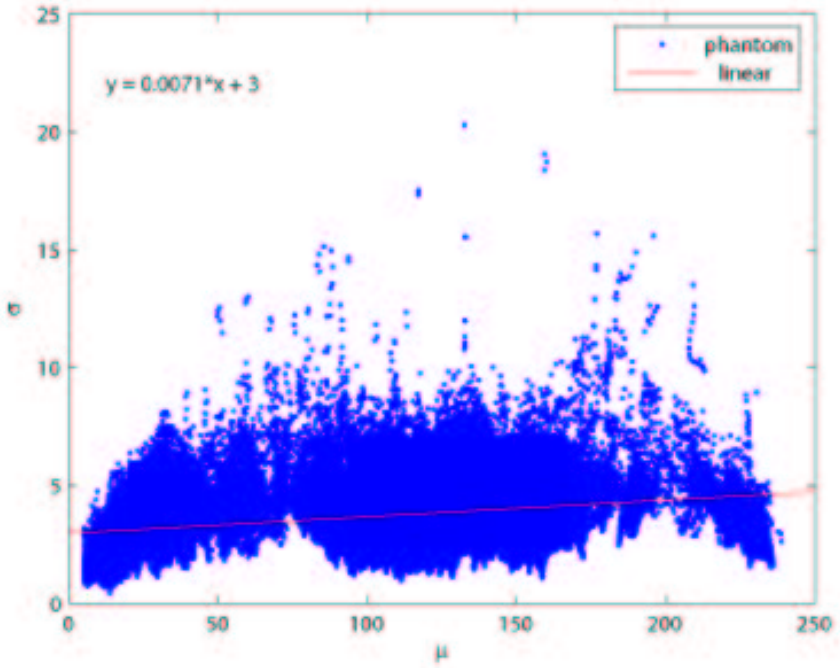
Figure 5.15 shows this dependency for phantom data. Notice that the variance on the standard deviation estimate is quite high, which is due to the fact that for some mean values there are very few samples.

Figure 5.16 shows the coefficient of variation ($\frac{\mu}{\sigma}$, see Section 4.2.2), which is intented to be constant throughout the range of mean values for other coherent imaging modalities. It is shown that it is not constant for log-compressed ultrasound images, thus not adequate as detector of uniform areas. These results suggest the adequacy of using only the local variance as uniform area detector and not the coefficient of variation.

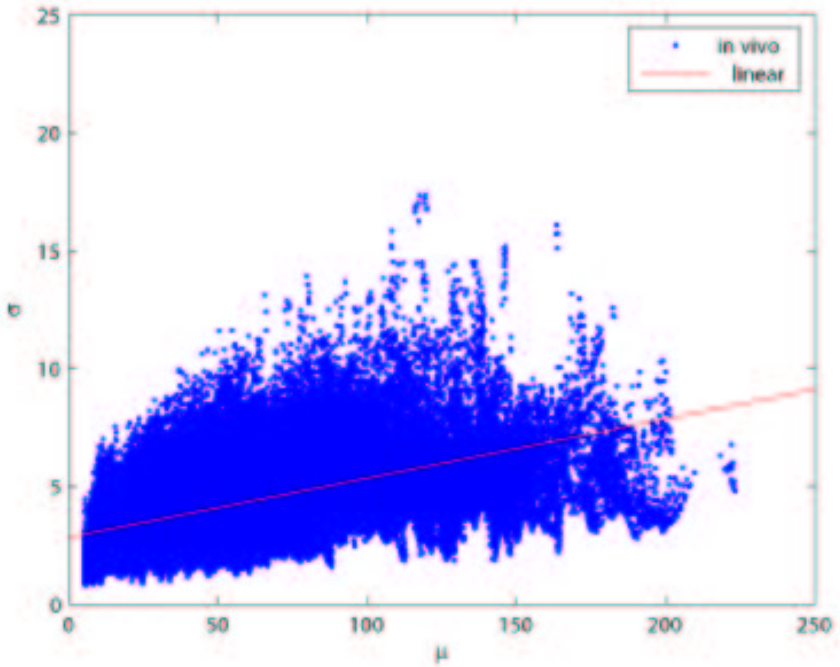
Autocovariance

Figure 5.17 shows the autocovariance profile in the lateral, axial and elevational directions, evaluated in uniform areas on the fruit phantom. Table 5.6 summarizes their numerical values. It can be observed that the autocovariance values follow the point spread function characterized in the previous section, that is, the biggest values are observed in the elevational direction, followed by the lateral and the axial directions.

The results presented in this section verify the simplified model of speckle in log-compressed ultrasound images proposed by [79] as colored Gaussian noise.

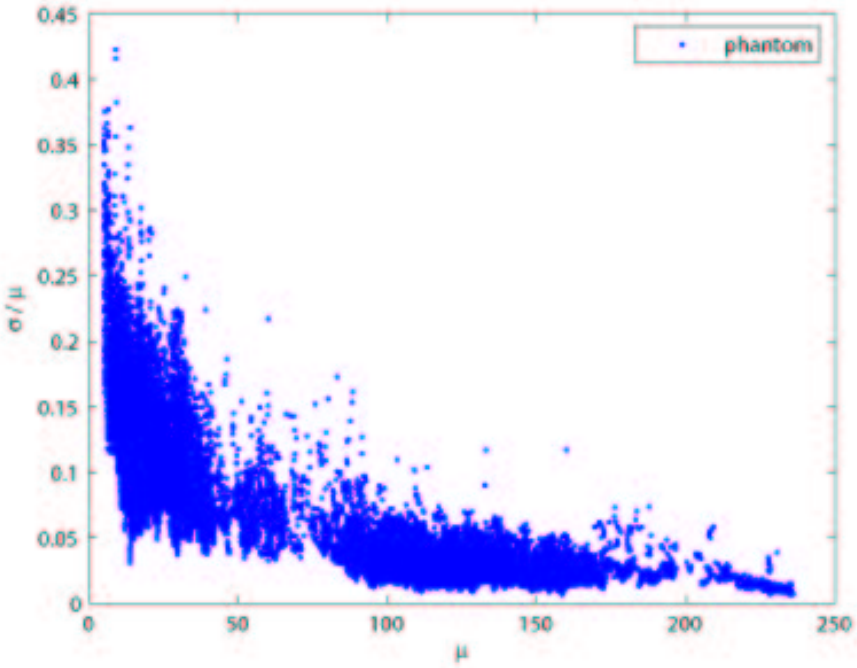


(*a*)

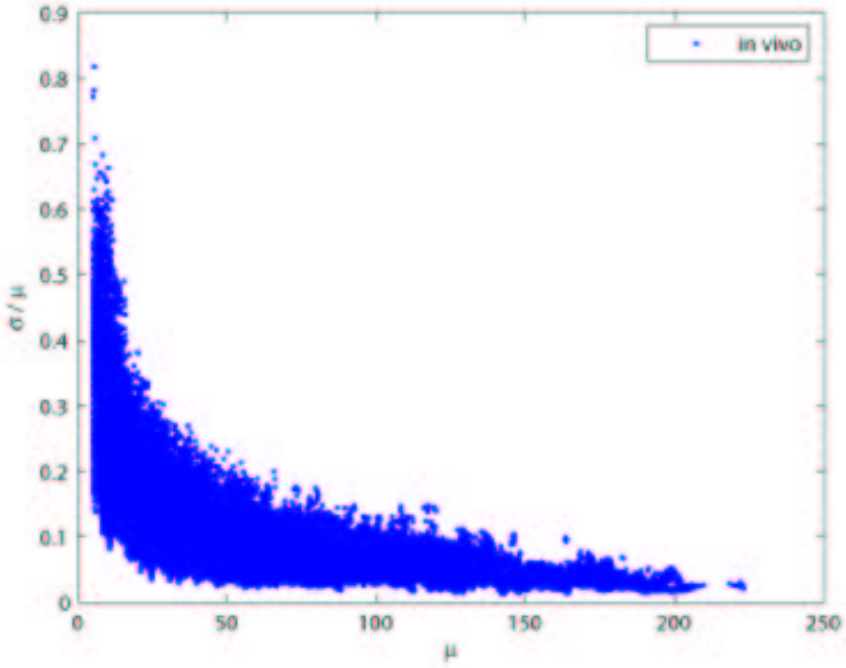


(*b*)

Figure 5.15: Relationship between mean and standard deviation on a) phantom data (mean of std = 3.8±1.1), b) in vivo data (mean of std = 4.3±1.7).

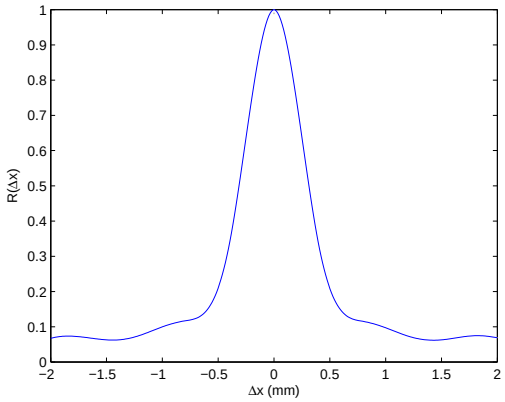


(*a*)

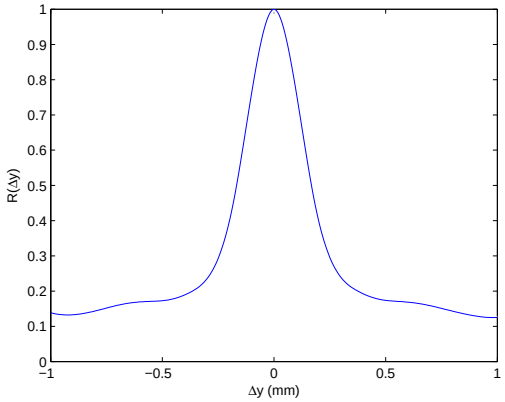


(*b*)

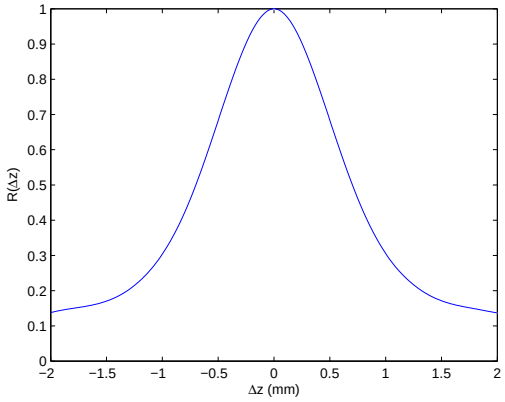
Figure 5.16: Coefficient of variation for log-compressed images on a) phantom data and b) *in vivo* data, showing it is not constant throughout all mean values.



(a)



(b)



(c)

Figure 5.17: Speckle autocorrelation profile for fruit phantom on uniform areas, in the a) lateral, b) axial and c) elevational directions. This metric shows the spatial statistical dependency.

	FWHM (mm)	σ	$2\sqrt{2\ln 2}\sigma$	ρ^2
lateral	0.59±0.29	0.29±0.21	0.68	0.9786
axial	0.30±0.09	0.12±0.04	0.28	0.9569
elevational	1.14±0.42	0.61±0.23	1.44	0.8076

Table 5.6: Estimation of the autocovariance of speckle, measured with the FHWM metric.

5.5 Registration

The different views need to be registered before combining them. Otherwise, the structures would appear at different positions, creating blurring artifacts. There are many factors which may contribute to the misalignment of the different views, such as mechanical misalignments, ultrasound imaging formation aberrations and patient movement during the scan.

Mechanical positioning errors may induce misalignment between the different views, due to uncalibrated offset and probe tilt between the different views. The mechanical precision of the linear encoders is of the order of 1 μm, much smaller than the resolution of the system. However, the position of the probe relative to the encoders needs to be calibrated in order to ensure the proper correspondence between the different acquisitions.

In second place, the formation of ultrasound images depends on the speed of sound (SoS), which indeed may not be constant. For this application, since all views are perpendicular to the breast tissue, a global misestimation from the true SoS is not significant since it will affect equally all views. Typical values of variation of SoS within breast tissue are small:

- Standard soft tissue, c_0 =1540 ms^{-1}
- Fat tissue, c =1450 ms^{-1} (- 6%)
- Muscle, c =1580 ms^{-1} (+ 3%)

The ratio of variation of the SoS is directly proportional to the produced misalignment, such as $\frac{c_0}{c}$. In other words, the worst case to misalignment would be about 3 mm at 5 cm of depth. However, this would be an extremely rare case where one view found only fat tissue and the others did not. In typical breast tissue, the expected variation between the different views is below the size of PSF.

In addition, tissue inhomogeneities may cause defocusing of the beamforming. This will not be corrected by the registration algorithm, and will have an effect of blurring the image, although it will be the same for all views.

At last, the movement of the patient can change the position of the breast tissue from their initial position. This movement can be due to breathing, cough, involuntary movement or repositioning in the scanning machine. We will develop a technique to compensate for such potential misalignment.

For all experiments, v_0 was chosen as the reference volume and the other views were registered against it.

5.5.1 Platform calibration

In order to calibrate the position of the probe with the mechanical encoders, a procedure was performed using the phantom "bubbles" (Section B.2), consisting of a set of individual reflectors. These reflectors should be at the same place in all acquisitions. The procedure consists in maximizing the cross correlation between the acquisitions in order to find the optimal offsets between the views. The search was done exhaustively in an interval of ± 0.4mm in lateral and elevational directions, with a step of ± 0.02mm. The offsets were found to be $\Delta x = 0.22$mm in the lateral dimension and $\Delta z = -0.18$mm in the elevational dimension, which correspond to approximately 1 pixel in each dimension. The same procedure was tested with the *in vivo* data, obtaining values which are close to those obtained with the phantom data, $\Delta x = 0.26$mm in the lateral dimension and $\Delta z = -0.24$mm. This could be useful in clinical routine if slight mechanical variations took place due to temperature changes or the like.

Table 5.7 summarizes the values of these experiments. The difference between offsets obtained on the phantom data or on the *in vivo* data where 0.04 mm and 0.06 mm in the lateral and elevational planes respectively, which is approximately 1/4 of pixel. In the axial dimension there was practically no difference.

	dx	dy	dz	Δx	Δy	Δz
phantom	0.22	0.07	-0.18			
in vivo	0.26	0.07	-0.24	0.04	0.001	0.06

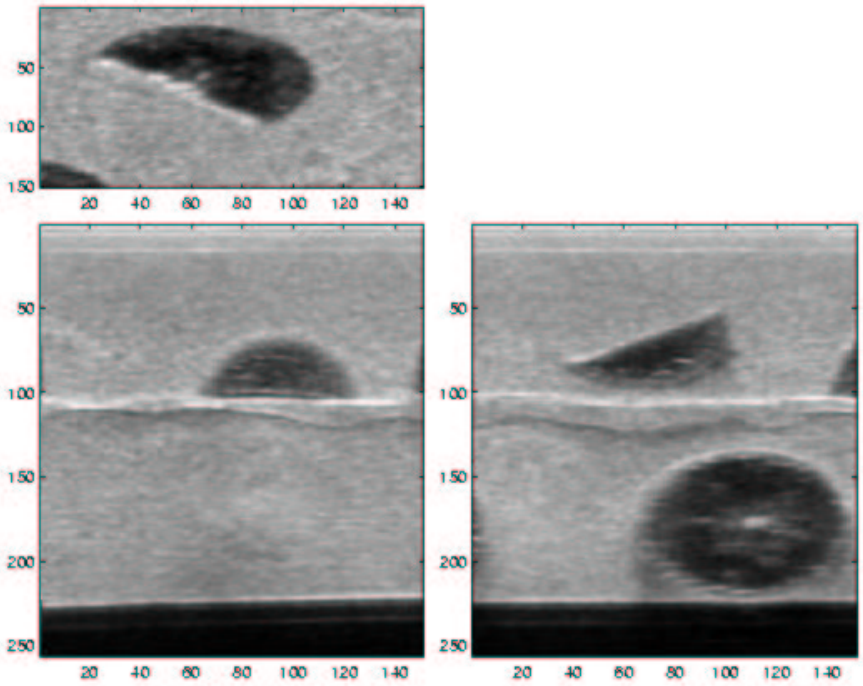
Table 5.7: Registration offsets (mm)

These values were evaluated along the axial dimension(y), showing no dependency ($\sigma = 0.006$ mm) on depth. This indicates that the probe tilt, which was manually calibrated, was correctly set to 90 degrees.

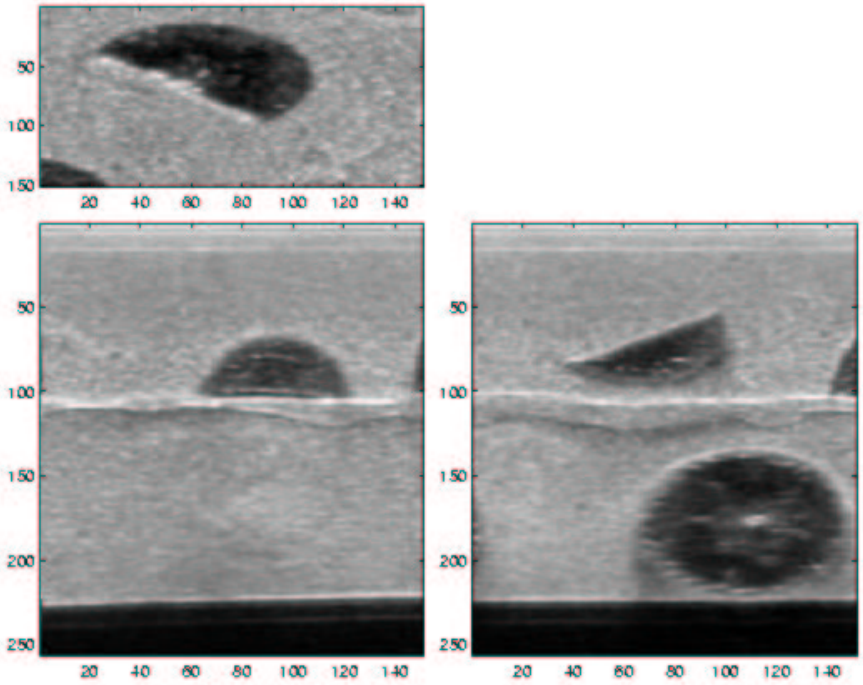
Figure 5.18 shows an example of offset registration on a fruit phantom, showing three slices (B-scan, E-scan, C-scan) of averaged volumes with and without offset compensation. Some structures, especially small fibers on the boundary of the fruit, are better delineated on the compensated data set.

5.5.2 *In vivo* data registration

In order to compensate both patient movements and deviations due to a speed of sound misestimation, a non-rigid registration procedure was applied. Rueckert *et al.* [131] reported the need of non-rigid registration for breast tissue on contrast-enhanced MR data, even with a coarser spatial resolution (about 1.5 mm). The method they proposed was based on a constrained free-form deformation (FFD)[98], using normalized mutual information as similarity metric [103]. For 3D breast ultrasound,



(*a*)



(*b*)

Figure 5.18: Average of four views: a) without offset compensation, and b) with offset compensation. Small structures are better preserved on the corrected data set.

for acquisitions with a linear array on a linear slicer obtained with different probe tilts, Krüker *et al.* [90] proposed a method of non-linear registration, based on: (i) performing a block matching using mutual information as similarity metric between the different views, (ii) selecting the blocks with the most significant shift and (iii) using the centers of these blocks as landmarks, transform the volume using a Thin-plate Spline interpolation [31]. On synthetically deformed data, this system improved deviations up to a mean deviation of 0.23 mm. On phantom measures, acquired with an aberrating layer to simulate beam deformations, and comparing the manually selected landmarks, the mean residual deviation was 0.30 mm.

Based on these works, we developed the technique that follows.

Similarity metric

We used a block-matching algorithm with sum of squared differences (SSD) as similarity metric. The acquisitions are obtained with the same machine, with the same imaging settings, therefore the relationship between the intensity levels is expected to be the identity. In this case, the SSD metric should work as well as any other metric, including the correlation coefficient or mutual information. Moreover, it has a lower computational cost, and the basin of attraction is typically larger [133], which permits a faster and more robust optimization of the criteria.

The size of the block is a critical parameter. Since the views are corrupted with speckle, the block has to be large enough for the matching to be robust with respect to speckle. Indeed, if no features are present in the block a noisy measurement will be obtained. Also, the point spread function is different in the different views, thus a perfect match between the different blocks cannot be expected. A possible approach to overcome this problem is proposed by Sroubek *et al.* [142], who perform registration and deconvolution at the same time, by estimating the point spread function along with its displacement. They also suppose a space-invariant point spread function. In our experiments, we used a size of 5x5x5 mm which turned out to be robust enough for the aforementioned problems. The search space was limited to ± 2 mm, and the total number of control points was a grid of 8x8x8mm (6x6x6 points).

Geometrical Transformation

To extend the displacements found by the block matching, we used a B-spline free-form deformation (FFD) transformation. The advantage of using FFD to other methods such as Thin-plate splines is that the deformation is locally controlled and they are computationally more advantageous. Indeed, breast tissue does not necessarily follow the implicit model of the Thin-plate spline transformation [131].

The FFD defines the transformation $\mathbf{T}(\mathbf{x})$, based on the deformation found by the block matching algorithm at each point of the uniform control grid. By denoting these control points as $\phi_{i,j,k}$, the transformation is defined for each point as [131]:

$$\mathbf{T}(\mathbf{x}) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(u) B_m(v) B_n(w) \phi_{i+l,j+m,k+n} \tag{5.3}$$

where $i = \lfloor x/n_x \rfloor - 1, j = \lfloor y/n_y \rfloor - 1, k = \lfloor z/n_z \rfloor - 1, u = x/n_x - \lfloor x/n_x \rfloor, v = y/n_y - \lfloor y/n_y \rfloor, w = z/n_z - \lfloor z/n_z \rfloor$ and where B_l represents the lth B-spline basis function [98]:

$$\begin{aligned} B_0(u) &= (1-u)^3/6 \\ B_1(u) &= (3u^3 - 6u^2 + 4)/6 \\ B_2(u) &= (-3u^3 + 3u^2 + 3u + 1)/6 \\ B_3(u) &= u^3/6 \end{aligned} \tag{5.4}$$

At each point $\mathbf{x}$, $\mathbf{T}(\mathbf{x})$ depends only on the closest four control points in each dimension. The points outside the domain of the image are considered to be fixed.

Registration Evaluation

Figure 5.19 shows the averaging of four views with and without the non-rigid transformation. Differences can be appreciated on the zoomed areas, where tissue delineation has been improved. It is expected that in clinical routine further potential displacements will occur, thus the need of this non-rigid registration step.

It is not obvious to evaluate the performance of elastic registration on *in vivo* data. Methods based on similarity metrics may be tendentious, since typically these metrics have been used to find the solution that makes them minimal. Other validation techniques, such as comparison based on manually selected landmarks are not applicable due to the very small displacements. The variability on manual clicks was found to be larger than the displacements. For ultrasound imaging of the breast, synthetic simulations with known elastic transforms, and measurements on phantom data measured with aberrated layers which simulate a non-linear deformation have prove to be quite effective in the literature [90].

To evaluate how large the displacements induced by the interpolated field $\mathbf{T}(\mathbf{x})$ are, we define the total mean displacement d as:

$$d = \frac{1}{N} \sum_{\forall \mathbf{x} \in \Omega} \|\mathbf{T}(\mathbf{x})\| \tag{5.5}$$

For each direction x, y, z, we compute the directional mean displacement d_u as:

$$d_u = \frac{1}{N} \sum_{\forall \mathbf{x} \in \Omega} \|\mathbf{T}(\mathbf{x}) \cdot \mathbf{u}\| \tag{5.6}$$

We compute the root mean squared error (RMSE) between the reference volume v_0 and the other volumes before and after the non-rigid registration. This gives us an idea of the overall similarity improvement obtained by the registration process. The RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{\forall \mathbf{x} \in \Omega} \left(v_0(\mathbf{x}) - v_i(\mathbf{T}(\mathbf{x}))\right)^2} \tag{5.7}$$

Table 5.8 summarizes these metrics for the different four views on *in vivo* data. The results show that the displacements are typically small, in the order of 200 μm for

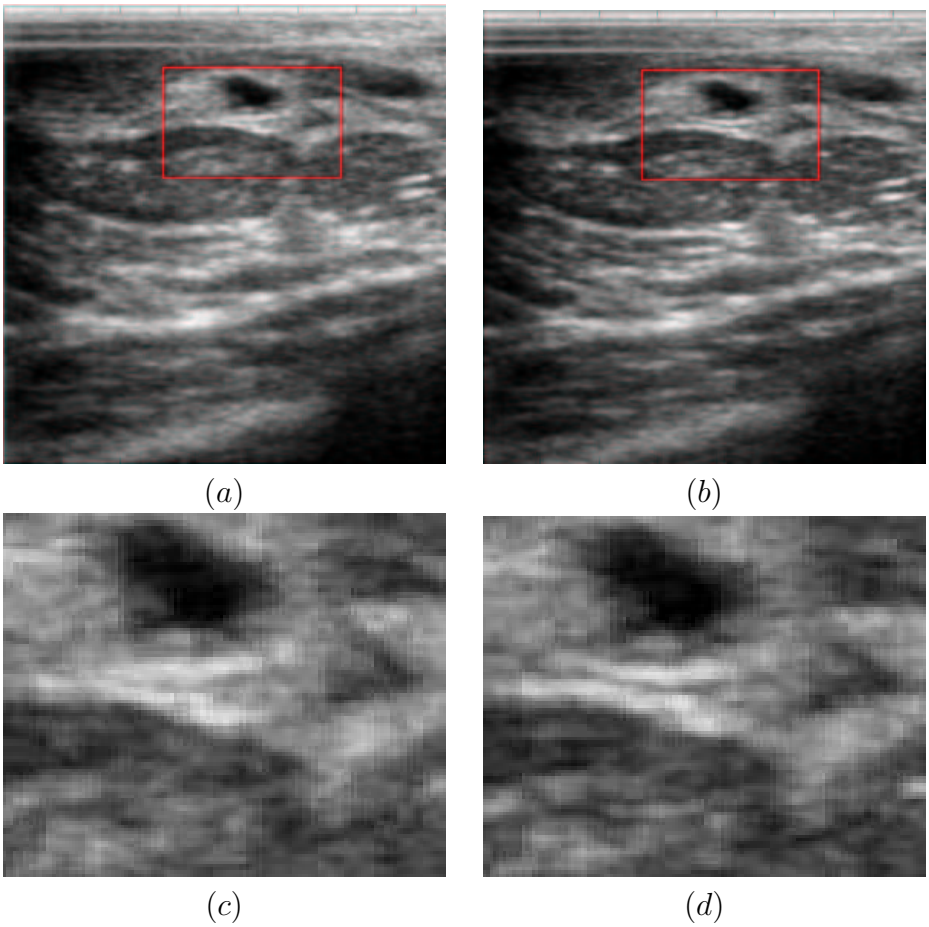


Figure 5.19: Example of non-rigid registration on *in vivo* data: a) average of views without registration, b) average of views with non-rigid registration, c) zoom on non registered view (a), d) zoom on registered view (b). For instance, the separation of the fibers below the breast mass can be better observed on the registered data set.

lateral and elevational directions, and in the order of 100 μm for the axial direction. These differences are equivalent, in average, to about one pixel. However, there are certain regions where the displacements are larger, up to 1 mm, thus the interest of using non-linear registration. This point is even confirmed by high variability obtained on the deformation field, which indicates that the field is not uniform. Larger deviations are also obtained in the elevational direction of each view, which may be due to the larger point spread function in this dimension. Overall, the RMSE between the reference image and the registered image is reduced by 1 level over 255 intensity levels, which is equivalent to 0.5%. This measure, while not quantitatively significant, shows that the overall similarity between the different views is improved by the registration process.

	$d(\mu\text{m})$	d_x	d_y	d_z	$RMSE_{reg}$	$RMSE_{ini}$	$\Delta RMSE$
v_{45}	215 ± 271	120 ± 192	51 ± 41	102 ± 232	19.63	20.42	-0.79
v_{90}	383 ± 223	169 ± 181	94 ± 69	289 ± 198	19.27	20.87	-1.60
v_{135}	419 ± 240	258 ± 271	68 ± 64	245 ± 159	20.14	21.17	-1.03
avg	339	182	71	212	19.68	20.82	-1.14

Table 5.8: Mean displacement (μm) and RMSE for *in vivo* data. The reference volume is v_0.

5.6 Restoration Results

In this section we present the results of the developed techniques on phantom and *in vivo* data acquired with the multiview strategy. Quantitative results on spatial and contrast resolution, critical parameters for medical diagnosis, are evaluated on the fruit and bubble phantom. Data acquired on patients are evaluated qualitatively.

5.6.1 Restoration Parameters

Table 5.9 gives an overview of the developed techniques to combine the different acquisitions. In this subsection we detail the computation of the parameters for each of the techniques according to the breast imaging application. For all techniques, M represents the number of views.

Noise power

Noise power parameter σ_n^2, or its standard deviation σ_n, is obtained from the image itself, as the mean of standard deviation in areas where the gradient magnitude is minimal. Local statistics are computed by convolving the signal with a Gaussian kernel. The size of this kernel should be large enough to contain the mean speckle size, in order to avoid noisy estimates within speckle homogeneous areas. The size of this kernel is estimated using the autocorrelation functions described in Section 5.4.2.

For case 1 of *in vivo* data, the autocorrelation values where $\sigma_s = (0.39, 0.21, 0.89)$ for lateral, axial and elevational directions, which corresponds to approximately 2, 3 and 4 pixels. Noise standard deviation was found to be 4.3 ± 1.1, and in order to avoid spurious artifacts an additional margin of three times the standard deviation of the estimate was used within the algorithm, giving final value of 7.6 gray levels.

Point spread function

The point spread functions h_i are assumed to be spatially variant Gaussian in all three directions, with the elevational (z) variance dependent on depth (y). These hypothesis were validated in section 5.4.1. The PSF of v_0 can be expressed as:

$$h_0(x, y, z) = G(0, \sigma_x) \cdot G(0, \sigma_y) \cdot G(0, \sigma_z(y) \tag{5.8}$$

Method	Equation	Section	Parameters
Average (avg)	$v_{avg}(\mathbf{x}) = \frac{1}{M}\sum_{i=1}^{M} v_i(\mathbf{x})$	4.2.1	-
Weighted Average (wa)	$v_{wa}(\mathbf{x}) = \sum_{i=1}^{M} \alpha_i(\mathbf{x}) v_i(\mathbf{x})$	4.2.2	α_i, σ_n, σ_s
Generalized Average (ga)	$v_{ga}(\mathbf{x}) = \left(\frac{1}{M}\sum_{i=1}^{M} v_i(\mathbf{x})^{\beta(\mathbf{x})}\right)^{1/\beta(\mathbf{x})}$	4.2.3	$\beta(\mathbf{x})$, σ_n, σ_s, γ
Maximum Frequency (mf)	$V_{mf}(\mathbf{f}) = max_i \{V_i(\mathbf{f})\}$	4.3.1	-
Wavelet Fusion (wf)	$\tilde{w}^l(\mathbf{y}) = \max_i \left\{\alpha_i(\mathbf{y}) w_i^l(\mathbf{y})\right\}$	4.3.2	α_i, σ_n, k
Multiview Deconvolution (md)	$v_{md} = \arg\min_{v_{md}} \left[(1-\lambda)\sum_{i=1}^{M} \|v_i - h_i * v_{md}\|^2 + \lambda\Psi(v_{md})\right]$	3	$\mathbf{h}$, σ_n, σ_s, λ

Table 5.9: Restoration Methods for 3D breast imaging

Given the nature of the acquisitions, the different point spread functions are rotated versions of h_0 at 45, 90 and 135 degrees in the lateral and elevational planes. The axial plane is the same for all acquisitions. We can express this relationship as:

$$h_\theta(x, y, z) = h_0(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta, z) \tag{5.9}$$

In practice, we divided the image in different slices of 5 mm of thickness and considered the PSF spatially invariant within this area, where the variances σ_x^2, σ_y^2 and σ_z^2 were estimated. The optimization of each of the parameters was performed with the golden search technique [51]. Figure 5.20 shows the estimation of the size of the PSF on *in vivo* data, compared with the one measured directly on the phantom data. It can be seen that overall there is a good match, although the size of the estimated PSF is always smaller for *in vivo* data. This may be due to two facts. First, the cross-channel relationship relies on the weak coprimeness of the different PSFs. In other words, if the PSF were strictly Gaussian, it would only measure the difference on their variances. And in second place, the size of bubbles of the phantom may also enlarge the size of the PSF measured directly on the phantom.

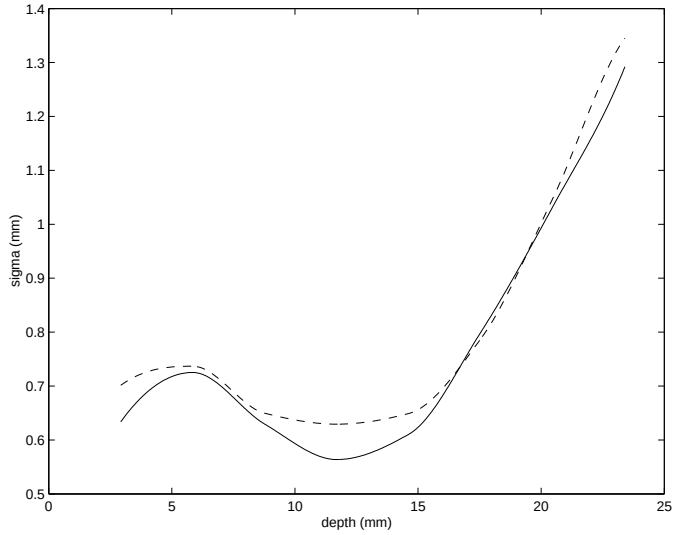


Figure 5.20: Elevational point spread function variance for bubble phantom (solid) and an in-vivo data set (dashed).

Regularization

The regularization term Ψ constrains the smoothness of the final solution. It is based on the norm of some function ψ applied over the neighbor difference operator Δ_i along

each dimension. In 3D, for a 6-neighborhood scheme it takes the form:

$$\begin{aligned}\Psi = &\sum_x \sum_y \sum_z \psi(\tilde{v}(x,y,z) - \tilde{v}(x-1,y,z)) + \\ &\sum_x \sum_y \sum_z \psi(\tilde{v}(x,y,z) - \tilde{v}(x,y-1,z)) + \\ &\sum_x \sum_y \sum_z \psi(\tilde{v}(x,y,z) - \tilde{v}(x,y,z-1))\end{aligned} \tag{5.10}$$

or in shorter notation,

$$\Psi = \|\psi(\Delta_x \tilde{v})\|_1 + \|\psi(\Delta_y \tilde{v})\|_1 + \|\psi(\Delta_z \tilde{v})\|_1 \tag{5.11}$$

where $\|\bullet\|_1$ represents the norm $\mathcal{L}1$.

We used a Lorentzian function [26], in order to smooth homogeneous areas while preserving edges. This function was first defined in the framework of anisotropic diffusion by Perona *et al.* [122]. Its definition can be found in Section 2.3.4.

The regularization parameter σ controls at which point the local averaging is stopped. In our case, we used the noise power as computed for the weighted averaging technique. Figure 5.21 shows the attractor of the regularization term alone ($\lambda = 1$), that is, without considering the deconvolution part for different values of the parameter σ. The hyperparameter λ determines the trade-off between edge restoration and image denoising. It was chosen manually to give equal weights to both data fidelity and regularization terms ($\lambda = 0.5$).

Iterative minimization of the energy was performed with the steepest descent approach. The used regularization term is not convex, and convergence is prone to be trapped into a local minimum. Therefore, it is important to properly initialize the estimated volume $\tilde{v}$. We have chosen to initialize the system with the average of the four views, which turns out to be robust and leads to satisfying results. Using other initializing volumes, such as a zero volume or one of the views, gives different but similar results. The iterative procedure is stopped when the variation in the solution falls below a certain threshold.

5.6.2 Phantom

Spatial resolution

Figure 5.22 shows the resulting PSFs obtained by combining the different views of the bubbles phantom, using the different presented techniques.

Table 5.10 reports the size of these PSFs as of the FWHM for a particular depth of 25 mm.

It can be appreciated that the average (Figure 5.22.b) of the different PSFs is much more isotropic than the PSF of a single acquisition, such as v_0. The size in the direction with coarse resolution, the elevational one, is greatly reduced by a factor of about three. However, the secondary lobes of the elevational plane from each view are present in the compounded point spread function. The weighted average (Figure 5.22.c) shows an identical behavior, since it is not intended to improve the spatial resolution. Generalized average (Figure 5.22.d) does improve the size of the

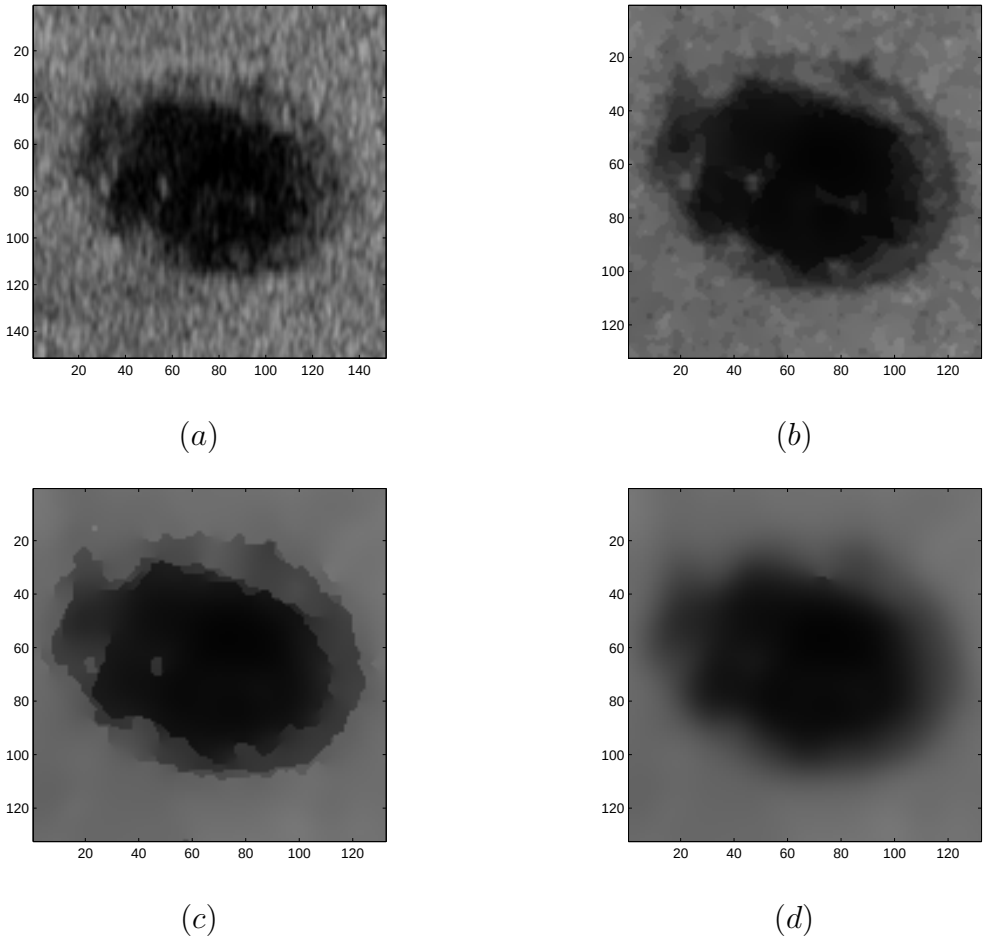
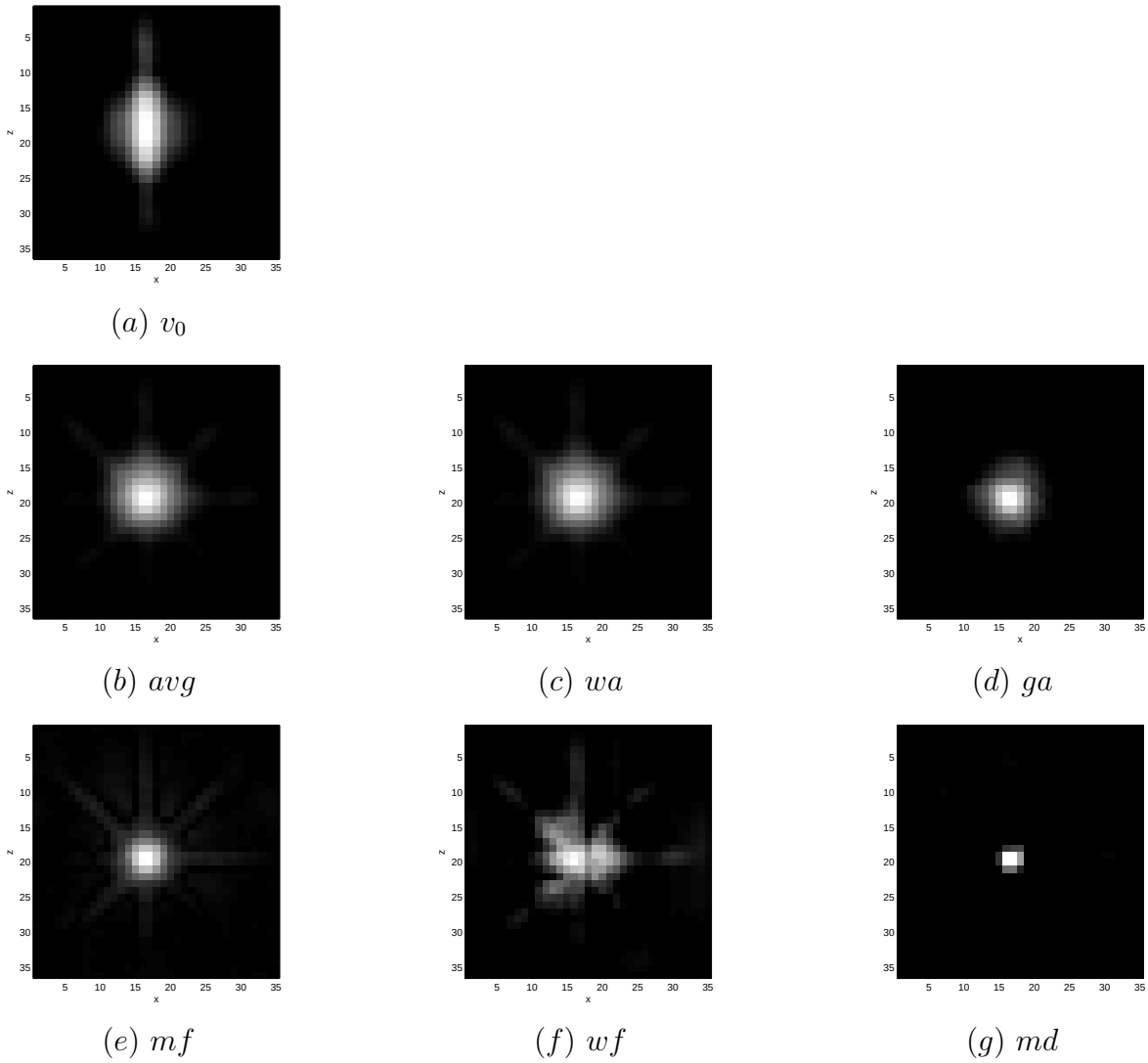


Figure 5.21: Assymptotic behavior of the regularization term ($\lambda = 1$) for different values of parameter σ: a) original view v_0, b) $\sigma = 3$, c) $\sigma = 7$, d) $\sigma = 15$.

FWHM (mm)	v_0	avg	wa	ga	mf	wf	md
lateral	1.10±0.3	1.47±0.2	1.72±0.2	0.95±0.2	1.08±0.3	2.17±1.0	0.75±0.1
axial	0.61±0.1	0.59±0.1	0.62±0.1	0.54±0.1	0.62±0.3	0.88±0.7	0.31±0.1
elevational	2.34±0.6	1.46±0.1	1.71±0.2	0.98±0.2	1.05±0.3	1.27±1.9	0.79±0.1

Table 5.10: Point spread function size (FWHM) at medium depth (y=25 mm) for lateral, axial and elevational directions on bubble phantom, for v_0 and methods: average (avg), weighted average (wa), generalized average (ga), maximum frequency (mf), wavelet fusion (wf) and multiview deconvolution (md). The anisotropic shape of the PSF is compensated by combining views from different angles.



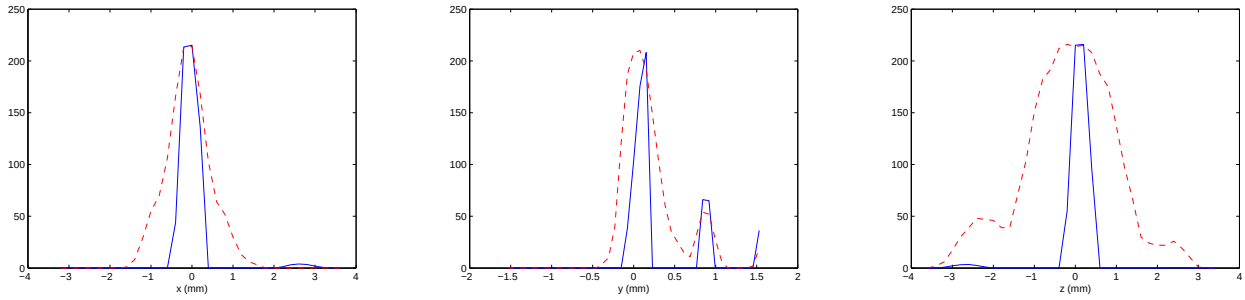
(a) v_0 (b) avg (c) wa (d) ga (e) mf (f) wf (g) md

Figure 5.22: C-scan slice of the point spread function, for a) v_0 and combined views: b) average, c) weighted average , d) generalized average, e) maximum frequency, f) wavelet fusion, g) multiview deconvolution (pixel size: 0.2x0.2mm).

point spread function, since in the case of complementary information it behaves as a minimum operator. Maximum frequency fusion technique (Figure 5.22.e) shows a reduced PSF kernel as well, although some artifacts that appear around the outer part of the central point. Wavelet fusion (Figure 5.22.f) does reduce the size of the PSF but some artifacts appear, creating a non-symmetric PSF. Finally, the blind multiview deconvolution (Figure 5.22.g) succeeds in reducing the size of the point spread function, although minor secondary lobes are still present.

Figure 5.23 shows the lateral, axial and elevational profiles of the multiview deconvolution method for the bubble shown in Figure 5.22. It can be appreciated that the most significant improvement of spatial resolution, with respect to a single acquisition,

is in the elevational dimension from 2.34 to 0.79 mm. Notice also that in the axial profile, clutter artifacts are a bit enhanced as if it were signal.



(a) (b) (c)

Figure 5.23: Point spread function profiles of combined views with the blind multiview deconvolution technique (solid) and original v_0 (dashed) for bubble centered at (x=10.40, y=26.39, z=5.40 mm). a) Lateral, b) axial, c) elevational profiles. There is a clear improvement of spatial resolution, especially in the elevational direction. Notice secondary bump in the axial profile due to clutter, which has been accentuated by the algorithm. Also, secondary lobes in the elevational profile have not been totally eliminated.

Contrast resolution

Noise power was evaluated on uniform areas on the fruit phantom. Table 5.11 describes the reduction of noise power for the different techniques. It can be seen how all techniques but maximum frequency significantly reduce the amount of noise. All spatial domain averaging (average, weighted average and generalized average) behave similarly (+2 dB), as expected in purely uniform areas corrupted with speckle. Wavelet fusion improves the noise power due to its de-noising stage, which acts as a local averaging (+3.62 dB). Similarly, the regularization term of the multiview deconvolution approach averages the voxel values with its neighbors, improving most notably the signal-to-noise ratio (+4.57 dB)

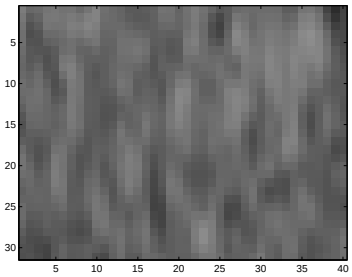
Figure 5.24 shows an uniform area on the original view and with the combined views. Figure 5.25 shows the intensity distribution of speckle in the original view v_0 and the combined one with the average, wavelet fusion and blind multiview deconvolution techniques. It can be appreciated how these techniques stretch the noise distribution, thus improving the contrast resolution.

5.6.3 *In vivo* data

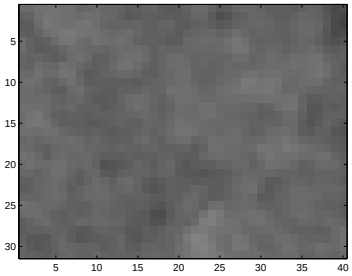
Exams on three benignant patients were done. Three slices of the original volume v_0 are shown in Figure 5.26. The first slice is a C-scan plane, where the anisotropic degradation due to the limited elevational resolution can be observed. Notice also how

	v_0	avg	wa	ga	mf	wf	md
std	3.78	3.08	3.09	3.10	4.71	2.59	2.25
ΔSNR	0	+2.09	+2.08	+2.04	-1.60	+3.62	+4.57

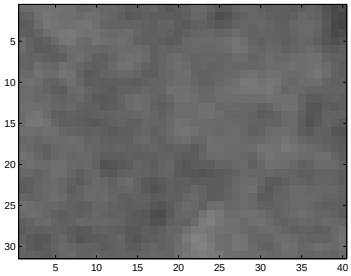
Table 5.11: Speckle standard deviation and ΔSNR (in dB) by combining different views with the following techniques: average (avg), weighted average (wa), generalized average (ga), maximum frequency (mf), wavelet fusion (wf), multiview deconvolution (md).



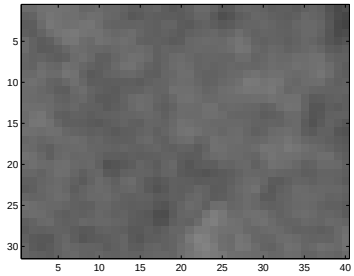
(a) v_0



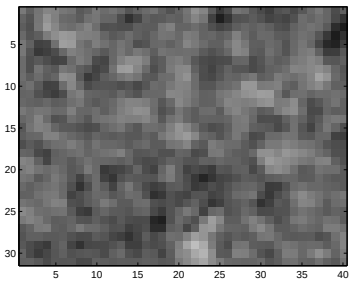
(b) avg



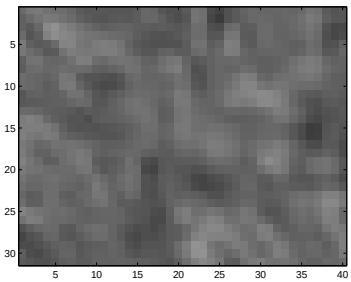
(c) wa



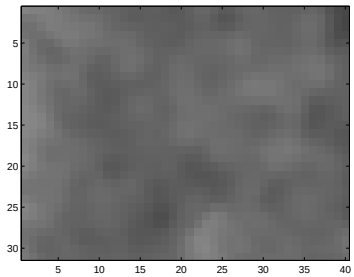
(d) ga



(e) mf



(f) wf



(g) md

Figure 5.24: C-scan of uniform area contained with speckle, for a) v_0 and combining methods: b) average (+2.09 dB), c) weighted average (+2.08 dB), d) generalized average (+2.04 dB), e) maximum frequency (-1.60 dB), f) wavelet fusion (+3.62 dB) and g) multiview deconvolution (+4.57 dB).

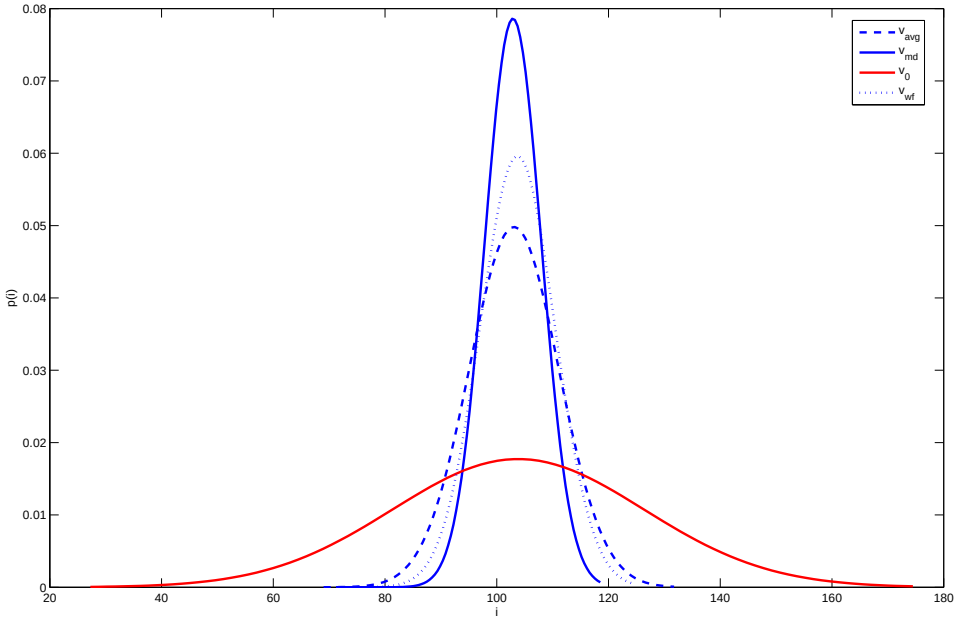


Figure 5.25: Speckle intensity distribution on uniform areas of Figure 5.24 for v_0 (red, solid), average (blue, dashed, +2.09 dB), wavelet fusion (blue, point line, +3.62 dB) and multiview deconvolution (blue, solid, +4.57 dB)

speckle patterning is present everywhere. The second slice (Figure 5.26.c) represents the B-scan, which corresponds to a standard two-dimensional ultrasound scan. This is the best resolution that can be obtained with this probe at the given frequency of operation. Notice how the size of speckle patterns is much smaller, showing a narrower distribution. Finally, the third slice (Figure 5.26.e) represents the E-scan plane, which is obtained by concatenating several B-scan planes. It can be seen that the resolution is lower, showing an overall smoothing aspect. Also speckle patterning is smoothed by the effect of the point spread function.

Figure 5.27 shows the C-scan of the combination of the four views v_0, v_{45}, v_{90} and v_{135}. The first image (Figure 5.27.a) shows the averaging of these views. Comparing to the original image v_0 (Figure 5.26.a), it can be observed that speckle has been notably reduced and that fibers are more clearly delineated. Most small structures, which in the original image are hard to distinguish between noise and signal, have disappeared, which indicates that those where not correlated within the different views. On the other hand, some fine details and small spots appear, which indicate their consistent presence in other views. Figure 5.27.b shows the results of using the weighted average technique. It can be observed that, in comparison to the averaged volume, the contrast of the structures has been enhanced, making little details more visible. This also helps distinguishing small fibers from the background. This is due to the fact that the technique preserves the views which show the highest signal-to-noise ratio of a

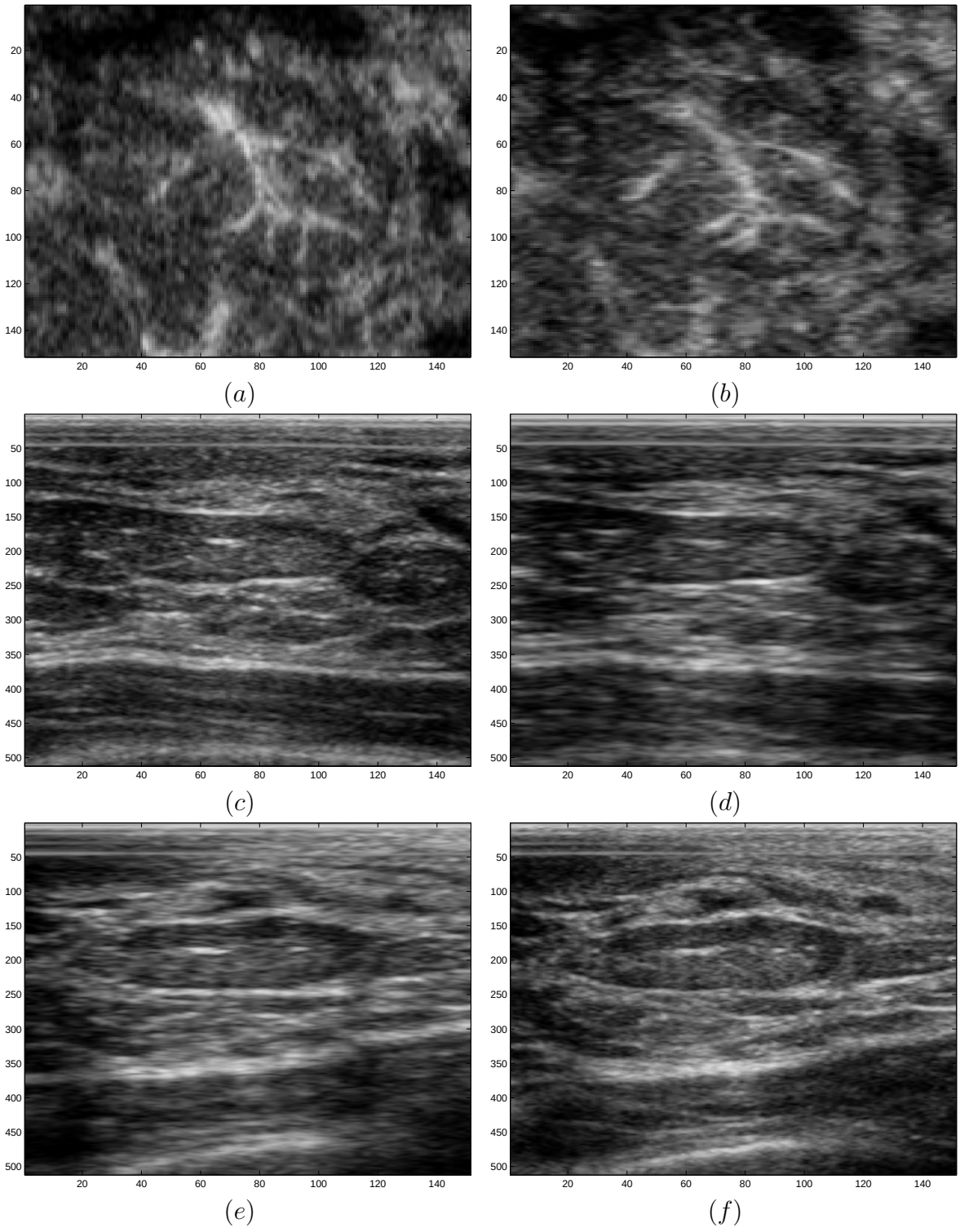


Figure 5.26: Original view of 3D breast tissue. Left, scanned at 0 degree (v_0). Right, scanned at 90 degree (v_{90}) . a, b) C-scan, c, d) B-scan, e, f) E-scan.

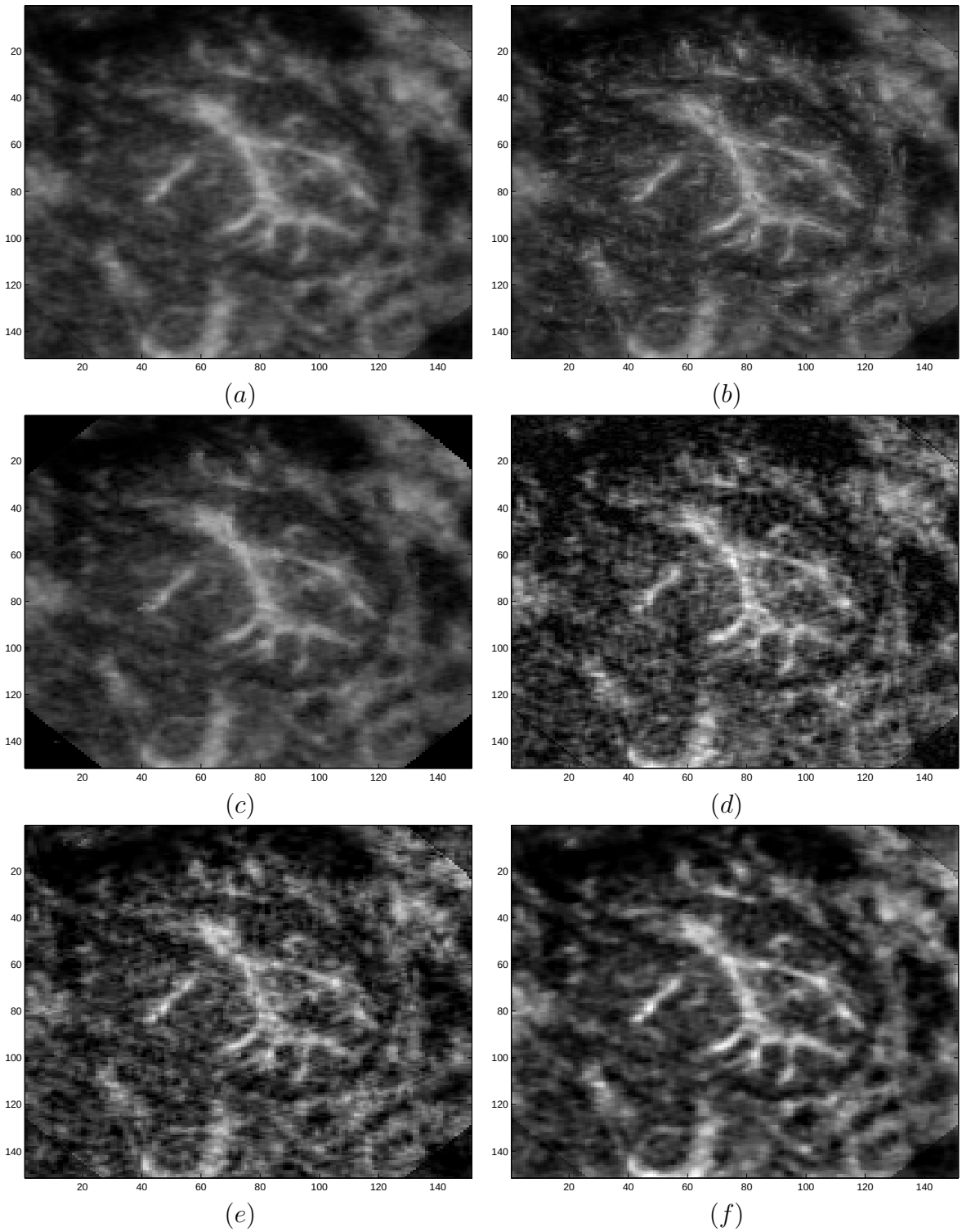


Figure 5.27: C-scan of combined views for *in vivo* data. a) Average, b) weighted average, c) generalized average, d) maximum frequency, e) wavelet fusion, f) multiview deconvolution.

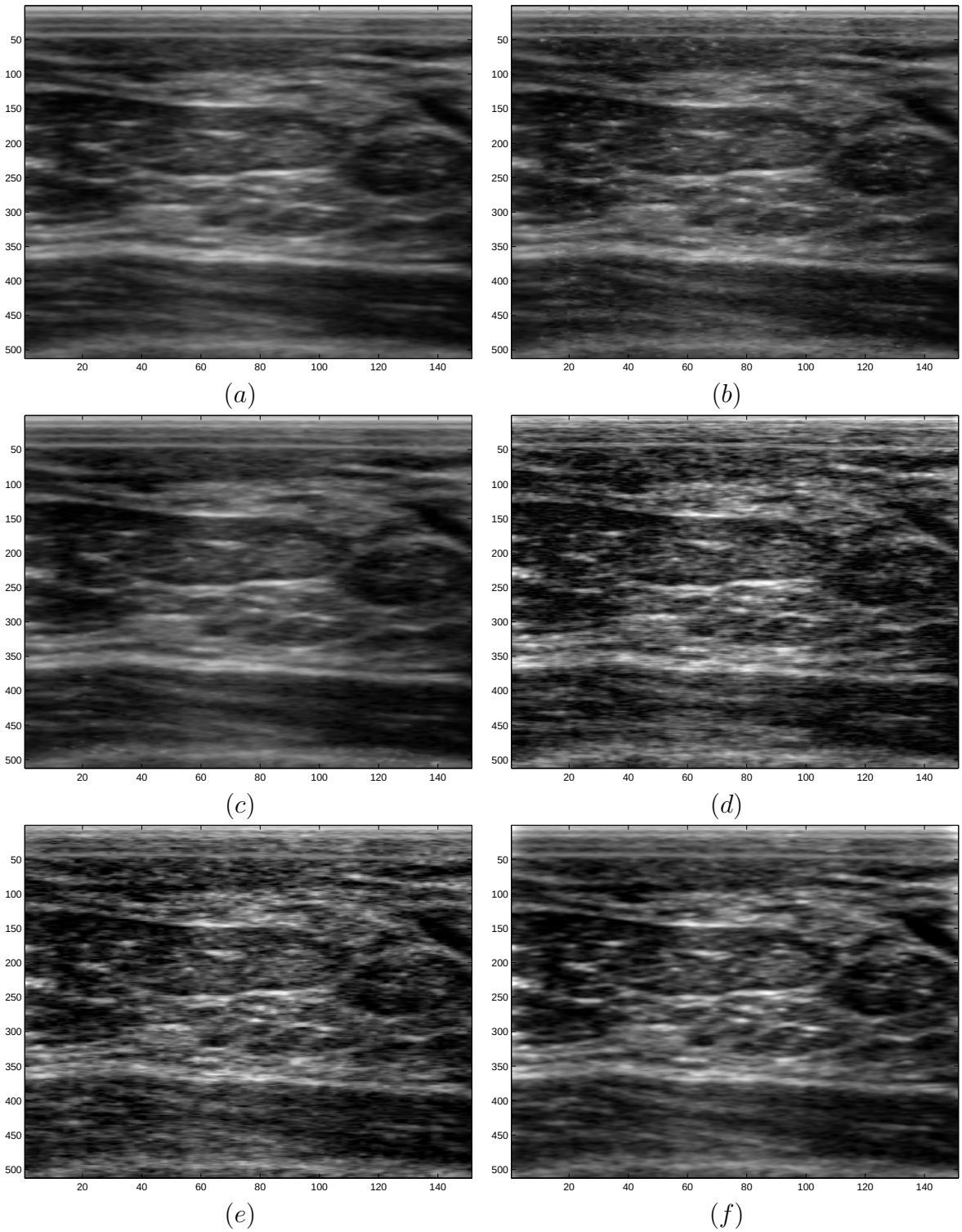


Figure 5.28: B-scan of combined views for *in vivo* data. a) Average, b) weighted average, c) generalized average, d) maximum frequency, e) wavelet fusion, f) multiview deconvolution.

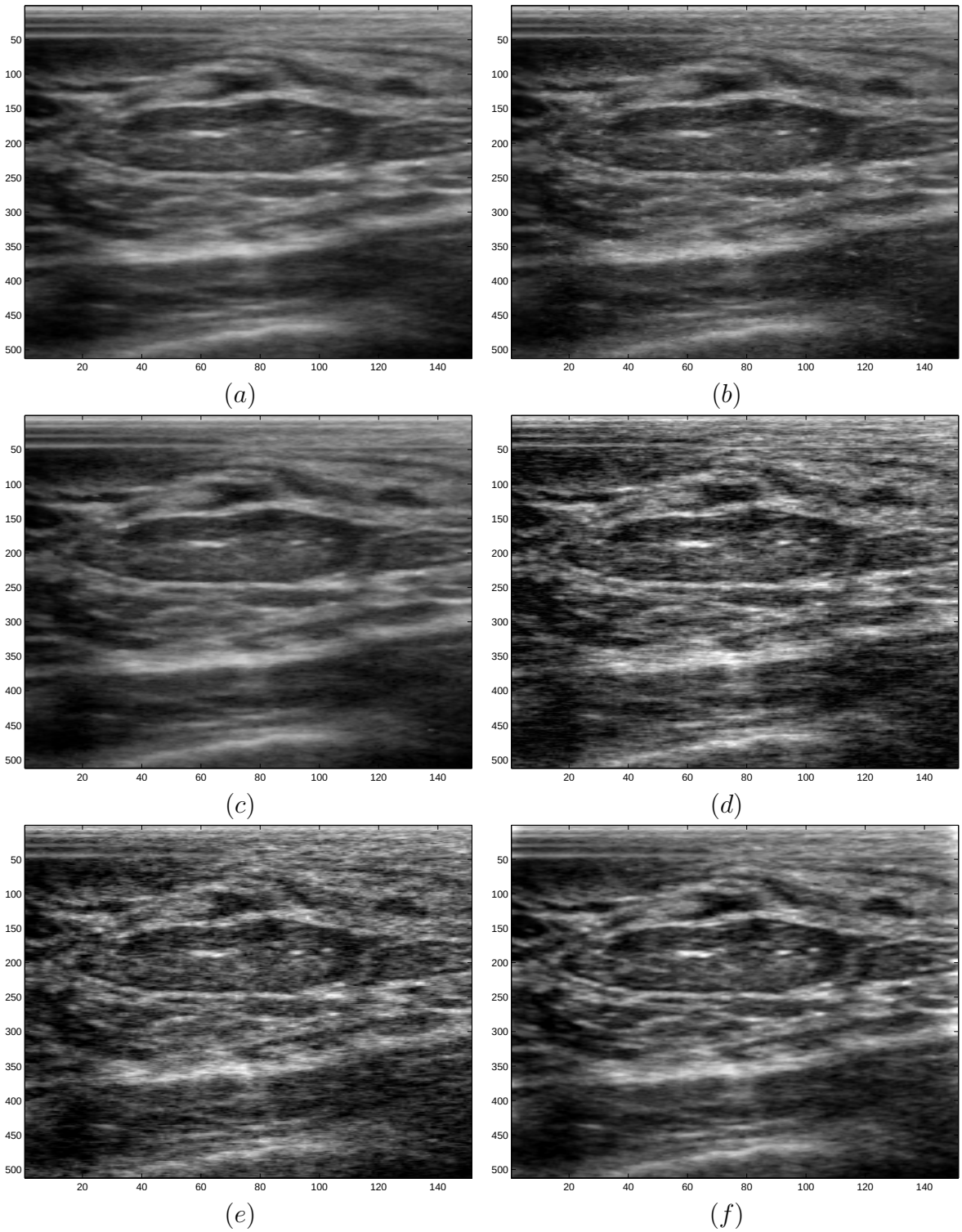


Figure 5.29: E-scan of combined views for *in vivo* data. a) Average, b) weighted average, c) generalized average, d) maximum frequency, e) wavelet fusion, f) multiview deconvolution.

given feature. The enhancement of small features appears irregular, which creates some visible artifacts throughout all the image. This is due to the fact that the saliency and speckle maps are a bit inexact at distinguishing small structures from speckle patterning. Figure 5.27.c shows the results for generalized averaging. Minor differences can be observed with respect to averaging, but for some specific areas at the edges of the fibers are better detailed. In the neighborhood of small spots, a halo effect appears due to the use of the maximum operator in this area. Although this ensures the presence of this kind of feature in the combined volume, it creates an artificial effect. Figure 5.27.d shows the results of maximum frequency technique. Speckle power has been increased in relationship to averaging, which makes difficult the evaluation of the image. However, all features present in the original views have been preserved, and the resulting volume is quite isotropic. Figure 5.27.e shows the results for the wavelet fusion approach. Similarly to maximum frequency approach, all features are present. However, the noise power is quite elevated, which induces to think that the denoising parameters have been underestimated. Some artifacts due to the thresholding of the wavelet coefficients are present, which could be accentuated if a wavelet coefficient threshold was augmented. Finally, Figure 5.27.f shows the results for the multiview deconvolution technique. Comparing to the average, it can be seen that a better contrast is obtained in structured elements, and that edges are better defined. Small fibers seem to be thickened by the effect of regularization.

Figure 5.28 shows the B-scan for the combined volumes. In comparison to the original view v_0, the averaging method has lost lots of details, but it has been greatly denoised. The reason is that the other views have a lower resolution in this plane. The weighted averaging technique does preserve some of these details, but many artifacts appear as well. Again, this is due to the difficult differentiation of small structures from speckle, needed within this algorithm. Generalized averaging does improve the contrast on some of the minor structures, although the overall effect is similar to averaging. The maximum frequency method is very similar to the original B-scan, which indicates that the best resolution is maintained. However, the noise power increase with respect to averaging is also visible in this plane. Wavelet fusion gives similar results, although tissues are better delineated in the smallest parts. Finally, multiview deconvolution gives the better resolved structures, showing all the details present in the original view but with a reduced speckle patterning. It must be noted that some of the spots present in the original B-scan were actually due to presence of spots in other parallel B-scan planes. Due to the effect of the point spread function, these appear in other B-scan planes, but are not necessarily present in the original image.

Figure 5.29 shows the E-scan for the combined volumes. In comparison to v_0, the averaged volume slightly improves the spatial resolution and reduces the speckle power. However, it does not reach the level of detail of v_{90} in this plane. Similar to the B-scan, weighted averaging shows better spatial resolution but at the cost of some artifacts. Generalized averaging behaves very similarly to averaging, showing a higher contrast on some fine structures and in areas where shadows are present. In this case, it behaves as a maximum operator that keeps any signal present. Maximum frequency

and wavelet fusion show a noisy behavior while the resolution is as good as the B-scan of v_0. Multiview deconvolution shows an increased contrast on small features and a good tissue delineation, although the level of noise looks higher than in the averaged view.

On the scanned patients there were no malignant cases, which does not permit to obtain solid conclusions on the utility of the system to discriminate malignant cases. However, the results obtained show that the critical parameters used for diagnosing such diseases have been improved, therefore it is reasonable to think it would make a significant difference with real cases.

5.7 Conclusions

We presented a system to obtain three-dimensional echography scans of the breast tissue. To improve important parameters for clinical diagnosis such as spatial and contrast resolution, we scanned the tissue with four different angles. In this way, the different views have differently orientated point spread functions and uncorrelated speckle patterning. By combining them, we obtain a volume which has an isotropic response, which means that we obtain the best resolution in every slice of the volume, contrary to a single acquisition, which has the best resolution only in the B-scan cut plane. Also, we obtained C-scan views that overcome the problem of the limited elevational resolution of 1D linear arrays.

To combine the different views, we used the techniques proposed in this thesis, namely: averaging, weighted averaging, generalized averaging, maximum frequency, wavelet fusion and multiview deconvolution. Multiview deconvolution performed best both quantitatively on phantom data and qualitatively on *in vivo* data. The size and anisotropy of the point spread function was greatly reduced with the deconvolution method, which improved the ability of image finner details of the system. In order to perform the deconvolution process, the point spread function is estimated using the cross-channel relationship developed in Chapter 3, which gave a good estimate according to measurements performed on phantom data. However, it tends to amplify artifacts which is not well modeled, such as clutter, which can create spiky artifacts that could mislead the radiologist. This can be solved by using more conservative parameters than the optimal ones as defined by SNR, for instance. The regularization term stabilizes the solution and reduces the speckle noise power in uniform areas. The needed parameters are estimated from the data itself, from the speckle intensity standard deviation and autocovariance. The deconvolution and regularization terms were equally weighted, although the trade-off parameter requires further study. The achieved improvements leaded to a better tissue delineation, clearer mass margins and a better mass contrast, which could potentially lead to better diagnosis. Pathologic cases were not present in the study with *in vivo* data, which does not permit to extract solid conclusions on the diagnosis abilities of the system.

Alternative techniques offered a good compromise between performance and simplicity. Wavelet fusion showed a good compromise between increase of spatial resolution and noise power reduction, although it exhibits some artifacts due to the non-

linear threshold process. This could be potentially overcome by the use of spatially adapting wavelets, such as curvelets [105]. Maximum frequency does preserve spatial resolution but does not reduce noise power. Averaging, on the contrary, does reduce noise power but does not preserve edges. It could be envisaged to combine these two techniques which showed complementary results and do not need any parameter. At last, weighted averaging and generalized averaging showed slight improvements over the averaging technique.

Here we presented the results with the parameters automatically estimated from data itself. Although these are close to optimal, in some cases better results could be obtained by manually tunning these parameters. This fact suggests that more advanced techniques to automatically find these parameters could be considered. Alternatively, advanced users could fine tune the parameters for a specific application (e.g. for a given ultrasound probe).

In general, the results were better for synthetic images or phantom images than for *in vivo* images. This is due to the fact the deviation of the real measurements from the used models, but also because in *in vivo* images all imaging needs are present at the same time; in other words, to better see a tiny fiber the algorithm has to preserve fine details and denoise at the same time at the same spot. This differs from the case in phantom data where we studied each factor at a time (i.e. point spread function size and speckle reduction). Prior to combining the different views, the different data sets were registered. The most contributing factor was found to be the calibration of the ultrasound probe with respect to the mechanical device. A non-rigid procedure to compensate patient movements within the scan was developed. Given the small size of the scanned tissue and the controlled environment, the detected movements were below the size of the point spread function, on the order of one pixel. However, it is expected that in clinical practice these movements could be increased and lead to blurring artifacts if not corrected.

We characterized the ultrasound data statistics of our system. We found that for a large range of intensity values, speckle intensity distribution is well fitted by a Gaussian distribution. The fit is worse for intensity values close to zero, where it fits better a Rayleigh distribution. Also, the correlation between the local signal mean and noise variance was found to be low, which validates the use of an additive noise model rather than a multiplicative one. The autocovariance of the noise was found to be related to the point spread function of the system, which validates the colored Gaussian model of speckle [79]. Finally, the point spread function of the system was also characterized, and found to be reasonably well fitted with a Gaussian curve of variance dependant on depth. Secondary lobes on the elevational plane violate this assumption, but experimental measures showed it was not relevant.

Future work lines may include integrating the deconvolution and registration process, since the registration process could potentially benefit of crisper image, as suggested by Sroubek *et al.* [142]. It could also include introducing the knowledge of the point spread function into the regularization term, in order to adapt better to the directions where there is more confidence on the data. Similarly, the knowledge of the point spread function could be introduced into the fusion techniques to better adapt to the

data characteristics. For instance, the maximum frequency technique would capture only the part of the spectra corresponding to each point spread function, averaging parts which are common, and similarly for the wavelet fusion or other decomposition (such as oriented laplacian pyramids) schemes.

Application-wise, the recent technological advances in planar arrays may also be benefitial for breast ultrasound imaging. In the next chapter we will deal with probes capable of obtaining real 3D data. However, given the size of the tissue, it looks like difficult that in the near future such kind of probe can be used without the help of mechanical platform to ensure a scanning of the whole tissue. Indeed, the spatial resolution of planar arrays so far is below of the linear ones. Therefore, the techniques developed in this chapter may also be applicable for these kind of probes.

CHAPTER 6

3D+T Cardiac Ultrasound Imaging

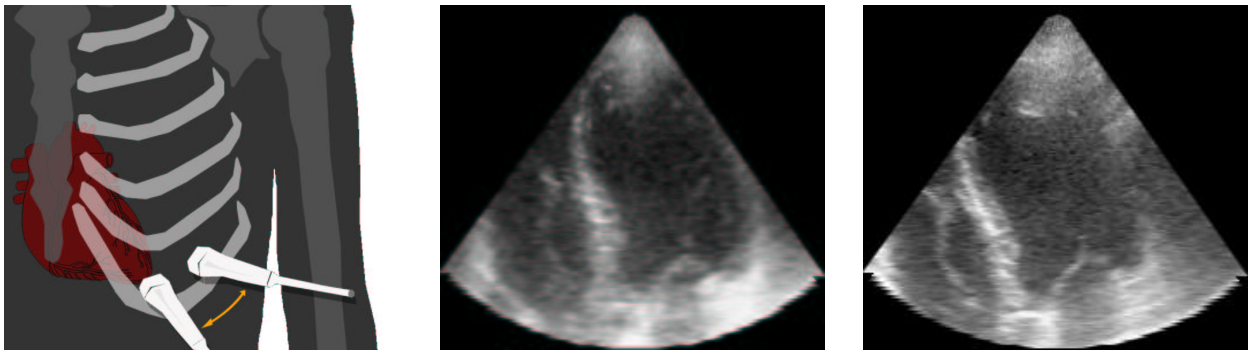
6.1 Introduction

Cardiovascular diseases (CVD) are the number one cause of death in the world. Only in Europe, they are responsible of nearly half of mortality, causing over 4.35 million deaths each year (as of 2005 [2]); on average, a person every 7 seconds dies from CVD in Europe. In the world, CVD are estimed to be responsible for one in every three deaths, affecting women and men of all geographic or economic conditions [7]. Many factors play a role in the development of CVD, such as smoking, diets, physical activity, diabetes or obesity. The two most common causes of death within CVD are coronary heart diseases and stroke. Medical imaging systems help the diagnosis of these diseases and empower doctors to conveniently deliver appropriate treatment to patients.

Ultrasound echography of the heart, also known as echocardiography, offers many advantages compared to other imaging modalities [104]. Principally, it is non-invasive, permits real-time acquisitions and it is relatively cost-effective. On the other hand, it suffers from some imaging limitations: (i) a restricted field of view, (ii) wall contrast dependency on insonifing angle, (iii) a coarse spatial resolution and (iv) a poor signal to noise ratio. Indeed, it is patient and operator dependent. Recently introduced into the market, real-time 3D (RT3D) echocardiography systems permit imaging tissue volumes, which enable new diagnostic applications, more accurate quantification methods and reduced time exam compared to 2D imaging (see Appendix A). On the other hand, it shows an even coarser resolution and a slower frame rate than standard echographic 2D imaging.

In this chapter we apply the multiview echography techniques to enhance imaging capabilities of 3D echocardiography systems [139]. Our goals are to (i) extend the field of view, (ii) capture features present in each view, (iii) improve the spatial resolution and (iv) improve the signal-to-noise ratio. Improving these image quality factors can lead to improved pathology diagnosis and more accurate quantification of cardiac function such as left ventricular volume. The proposed method consists in fusing several views from different acoustic windows. Acoustic windows refer to the external part on the body from where the heart can be imaged, because ultrasound cannot propagate

through lungs and the rib cage. In particular, we will focus on acquisitions from displaced apical windows. Figure 6.1 shows an example of two acquisitions from different apical acquisitions. The proposed method consists in registering these different views and fusing their intensity values into a single volume.



(*a*) (*b*) (*c*)

Figure 6.1: a) Schema of two different apical window probe positioning, b) slice of RT3D standard apical view, c) slice of RT3D displaced apical view. The goal of our method is to combine views b) and c) to increase the field of view, improve wall contrast and attenuate artifacts.

The combination of echocardiographic acquisitions from different acoustic windows has been addressed in the literature for linear arrays. Legget *et al.* [99] used a free-hand 2D system, consisting in a linear probe with a magnetic-positioning sensor. The LV surface was reconstructed using features extracted from the different views. Ye *et al.* [165] used a rotational 3D probe (a 2D probe that acquires slices at fixed angle increments), which on the one hand gives a higher resolution image at each slice but on the other hand generates a sparse volume of data that is prone to temporal artifacts. To register the views, an external optical sensor tracked the position of the probe. Features were extracted by a phased-method based algorithm called Feature Asymmetry [114] and was used to refine the registration. The fused image is constructed by combining these features, weighted by the viewing geometry [100]. Recently, Grau *et al.* [59] tackled the problem for RT3D echocardiography, using an estimate of phase and orientation to define the weights or the fusion step, but no reference is made on how images are registered. The significant overlap of the different volumes enables their registration without the need of an external positioning probe.

This chapter is divided as follows. We start introducing the imaging needs from the clinical perspective in Section 6.2, in order to understand which are the imaging parameters of importance. Data acquisition materials and protocols are described in Section 6.3. Registration techniques, which are needed to put in correspondence the different views, are described in Section 6.4. Results of combining those acquisitions, are given in Section 6.5. And finally, conclusions are drawn in Section 6.6.

6.2 Clinical Needs

6.2.1 The Heart [3]

There are two sides of the heart. The right side of the heart pumps deoxygenated blood from the body veins to the lungs, where the blood picks up oxygen. Then the oxygenated blood returns to the left side of the heart, which pumps the blood out to the body arteries. There are five main areas in and around the heart :

- The *atria*, which collect the blood returning to the heart. The right atrium (RA) collects the blood from the body, and the left atrium (LA) from the lungs.
- The *ventricles*, which receive the blood from the atria and pump the blood out of the heart; the right ventricle (RV) to the lungs and the left ventricle (LV) the circulation system.
- The four *valves* of the heart. These valves allow blood to move only in the proper direction in the heart without allowing any blood to leak back in the other direction. These four valves are the tricuspid valve (TV), between the right atrium and ventricle; the pulmonary valve (PV), between the right ventricle and the pulmonary artery to the lungs; the mitral valve (MV), between the left atrium and the left ventricle; and the aortic valve (AV) between the left ventricle and the aorta to the body.
- The *veins* that carry the blood back to the heart from the body and the lungs.
- The *arteries*, the aorta (which carries oxygen-rich blood from the left ventricle to the rest of the body) and the pulmonary artery (which carries oxygen-poor blood from the right ventricle to the lungs).

Figure 6.2 shows an illustration of the heart, with the aforementioned areas, and Figure 6.3 shows how the heart is seen in a echocardiography exam for an apical view. Table 6.1 summarizes the abbreviations used throughout this chapter.

RA	right atrium
LA	left atrium
RV	right ventricle
LV	left ventricle
AO	aorta
PA	pulmonary artery
MV	mitral valve
TV	tricuspid valve

Table 6.1: Common abbreviations of heart parts.

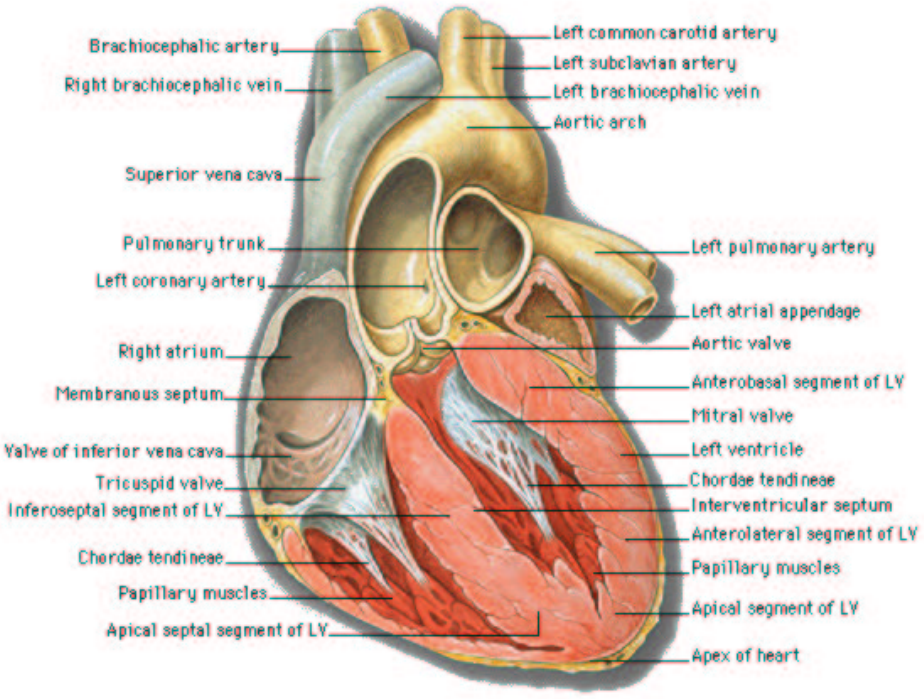


Figure 6.2: Heart Anatomy [8].

6.2.2 Common diseases [3]

We give a brief overview on some of the most common diseases detectable with echocardiography, and particularly, with dilated hearts, which is one of the main applications of the presented technique. The complete list of heart diseases is out of the scope of this document.

Arrhythmias

An arrhythmia is an abnormal heartbeat that may be unusually fast (tachycardia) or unusually slow (bradycardia). It may be related to a previous heart condition (e.g., previous damage from a heart attack) or to other factors (e.g., caffeine, stress, medications). Diagnosing an arrhythmia is very important, because the longer an arrhythmia lasts without detection or treatment, the greater the chances of permanent damage and additional heart dysfunction. Severe arrhythmias may lead to fainting or even sudden cardiac death.

Cardiomyopathy

Cardiomyopathy is a type of heart disease in which the heart becomes abnormally enlarged (enlarged heart), thickened and/or stiffened. As a result, the heart muscle's ability to pump blood is usually weakened. This condition is generally progressive and may lead to heart failure. Cardiomyopathies cause symptoms including shortness of

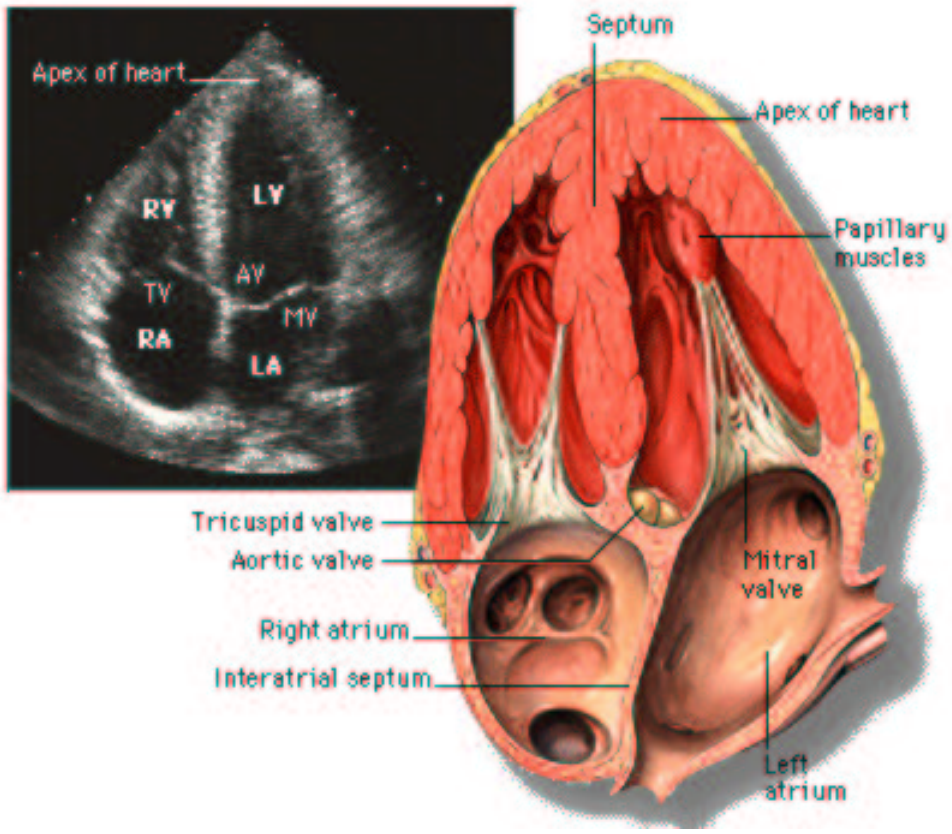


Figure 6.3: Heart Anatomy as seen in a 2D apical echocardiographic exam [8]

breath, chest pain, fainting, dizziness and a reduced ability to exercise. The muscle damage that develops with all types of nonischemic cardiomyopathies can lead to arrhythmias.

Cardiomyopathies may be caused by a wide range of conditions, including chronic diseases, alcoholism, viral diseases, heart attacks and many others. An affected heart may grow larger either by dilatation, thickening hypertrophy or both. Additionally, the heart may suffer from a reduced ability to relax. Abnormalities found in cardiomyopathy include:

- Thickened and/or dilated ventricles, especially the left ventricle. The upper chambers (atria) may also be involved and enlarged.
- Scar tissue, possibly left over after a heart attack.
- Overall enlargement of the heart.

Cardiomyopathies are often considered ischemic (resulting from a lack of oxygen) and nonischemic.

Ischemic cardiomyopathy is a chronic disorder caused by either recurrent heart attacks or coronary artery disease (CAD) - a disease in which there is hardening

(atherosclerosis) of the arteries on the surface of the heart. CAD often leads to episodes of cardiac ischemia, in which the heart muscle is not receiving enough oxygen-rich blood. Additionally, as a result of one or more large heart attacks, the heart enlarges, reducing the heart muscle pumping function.

Nonischemic cardiomyopathies are less common, progressive diseases. Unlike ischemic cardiomyopathies, which tend to develop in older adults, nonischemic cardiomyopathies frequently occur in young people. Nonischemic cardiomyopathies can be difficult to diagnose because many are idiopathic (i.e., their cause is unknown). However, known causes include genetic factors, viral infection, build-up of fat and proteins (amyloidosis) in the heart muscle, or an excess of iron (hemochromatosis) in organs such as the heart. The excessive use of alcohol or other substances can also play a role in the development of the disease.

There are three main types of nonischemic cardiomyopathies:

- Dilated cardiomyopathy (including peripartum cardiomyopathy and alcoholic cardiomyopathy), which involves dilation or enlargement of the heart's ventricles and is usually accompanied by an increase in cardiac mass. This often affects young people. Although ischemic cardiomyopathy physically resembles dilated cardiomyopathy, many physicians separate ischemic cardiomyopathy caused by CAD into its own class.

- Hypertrophic cardiomyopathy, which involves an abnormal growth of muscle fibers in the heart muscle, usually in the left ventricle. In this case, the volume of the left ventricle is normal or reduced, but the additional muscle fibers prevent the chamber from relaxing completely after contraction (diastole), making it a diastolic dysfunction. This is usually considered a genetic disorder.

- Restrictive cardiomyopathy, which means the heart muscle cannot adequately relax after contraction, making it unable to fill completely with blood. This condition is distinguished from some forms of hypertrophic cardiomyopathy because the left ventricle is frequently normal sized. This cardiomyopathy is more common in the tropics that other forms of cardiomyopathy. Less common types of nonischemic cardiomyopathy include arrhythmogenic right ventricular dysplasia (ARVD).

Other

Valvular heart disease is the name given to any dysfunction or abnormality of one or more of the heart's four valves, including the mitral valve and aortic valve on the left side, and the tricuspid valve and pulmonic valve on the right side. In a normally functioning heart, the four valves (flaps made of tissue) keep blood flowing in one direction and only at the right time. They act as gates that swing open to allow blood to flow through and then tightly shut until the next cycle begins.

Congenital heart disease (CHD) is a heart-related problem that is present since birth. There are many different kinds of specific congenital heart defects. They may

affect various heart structures, such as the valves, the veins leading to the heart, the arteries leaving the heart or the connections among these various parts.

6.2.3 Imaging Goals

Ultrasound imaging has been widely used to diagnose heart pathologies. It permits the evaluation of many of the parameters of interest, including: (i) the size of the chambers, including their volume and the thickness of the walls, (ii) the local pumping function, measured typically by assement of local displacement and deformation of the cardiac muscle, (iii) wall synchronicity, measured by analyzing the movement of each of the wall segments, or (iv) valve function. These parameters are critical to diagnose the aforementioned heart pathologies, such as cardiomyopathies. Echocardiology offers a simple way to obtain quantitative measurements for those parameters. The objectives of our method are to improve the imaging capabilities of echocardiographic systems in order to help doctors to better evaluate the mentioned diseases.

Field of View

In order to properly evaluate the synchronization of all parts of the heart, it would be desirable to have into a single view the four heart cavities and the main walls. However, a limited field of view is obtained with transthoracic acquisitions. This is because ultrasound waves cannot penetrate hard tissues, such as bones or ribs, nor lungs, creating shadows on the image. Therefore, the size of the probe, which must be small enough to fit in between the ribs, limits the field of view. Ultrasonographers "search" through the different acoustic windows in order to capture the needed information for their diagnostic. Typically, in a standard apical view, named because the probe is placed on the chest close to the apex of the heart, the apex and the right ventricle are difficult to see at the same time. By placing the probe sightly to the sides, the apex and the right ventricle will be visible, but other parts of the left ventricle might be cut off. By fusing the different views, an extended field of view can be obtained. This is of critical importance in certain cardiomyopathies were the heart is dilated.

Wall Contrast

Tissue contrast depends on the insonification angle. If tissue interface and ultrasound beam are parallel, weak echo amplitudes will result in poor contrast. This effect is specially noticeable when insonifing through the different acoustic windows, since the insonification angle largely vary between the views. Also within each acoustic window some walls are more visible than others due to this effect. In fact, it requires a great expertise to the ultrasonographer to find an appropriate angle. Even then, not all interesting features are captured in a single view. Besides the simple fact of not seeing the walls, this actually implies that quantification tools, such as automatically detecting the contours of the left ventricle, will not have enough information to succeed.

If a wall or a part of a wall is missing, besides not being able to evaluate its synchronicity, many other parameters might be difficult to obtain. for instance, to cal-

culate the ventricular volume, the doctor or the automatic procedure has to "guess" where the wall is, yielding into a lack of accuracy. Indeed, many automatic algorithms may simply fail to obtain a reliable shape of the ventricle if walls are not completely visible. Actually, automatic segmentation of LV is still an open problem, which typically does work reasonably well for images of good to medium quality. If all walls are present, presumably the performance of those algorithms will improve.

In order to evaluate certain cardiomyopathies, it is important to measure the thickness of the endocardium. The contrast of the endocardium may be quite low in some of the patients, being difficult to distinguish from the interior of the cavity. This low contrast may be improved by superimposing the different views, thus better evaluating the local function.

Spatial Resolution

The need for high frame rates does not permit the use of sophisticated beamforming techniques which would improve spatial resolution. Instead, plane wave techniques are used and beamforming is performed only at reception, resulting in a coarse spatial resolution. Indeed, spatial resolution is irregularly distributed, getting sparser as the distance to the focal points increases.

Accuracy of LV quantification also depends on tissue delineation, as blurred borders may induce false measurements. Views with different angles have a different PSF, and the spatial characterization diversity may be used to obtain better delineated tissues.

Signal-to-noise Ratio

As with any coherent imaging system, ultrasound echographic images are degraded by speckle noise (see Sec. A.3.2). This may obscure some small features, and also make harder the task of automatic quantification algorithms, such as left ventricle segmentation. If the spatial diversity is large enough, the views will have uncorrelated speckle, enabling denoising of the imaged tissues.

Patient and Practitioner Independency

Factors such as smoking or obesity heavily impact the echogenecity of the patient, making the diagnosis harder. When transthoracic acquisition do not provide sufficient image quality, transesophageal acquisitions are performed, where the echographic transducer is placed within the esophagus. Currently, there are no transesophageal planar arrays, so this imaging modality is limited to 2D, so no real-time 3D imaging is possible on this modality.

Ultrasonographer skills also play an important role in the quality of the echography image. It would be desirable to improve the image quality even in the poor echogenic cases. Potentially, from several low-quality exams a better image quality exam could be reconstructed.

6.3 Data Acquisition

6.3.1 Materials

RT3D data were acquired with a Philips Sonos 7500 system with Live3D Echo, using a x3-1 matrix array. Each acquisition is completed within four consecutive cardiac cycles, triggered by the electrocardiogram (ECG). The frame rate was of 25 ms, which implies that the total number of frames per cycle depends on the patient's beating rate. Indeed, this may vary between different cardiac cycles, e.g. due to variations on breathing. Typically, there are between 15 and 25 frames per cardiac cycle.

Data were resampled from the original line data to Cartesian coordinates through a scan conversion process. The final size of the volume depends on the depth settings used, and a typical volume is of 136x130x123 mm^3 in lateral, elevational and axial dimensions respectively. The pixel size depends on the scan conversion process, with typical size of of 0.6x0.6x0.5 mm^3 in lateral, elevational and axial dimensions respectively.

No external position-encoder device was used.

6.3.2 Method

The acoustic windows used were a standard apical (approximately between the fourth and fifth rib, about 5 cm to the esternon), and two displaced apical by approximately 2 centimeters to the left and right within the same intercostals window. One of the planes of the 3D volume was coincident with a standard 2D apical four-chamber view. Patients were asked not to breath in order to avoid motion artifacts during the acquisition, and to obtain a regular cardiac cycle. Other regular acquisitions through standard acoustic windows, such as parasternal and subcostal, were also performed.

The task of registering the images is easier if the acquisition process follows some rules. Gain should be set so that image contrast is regular throughout the image. Special attention should be paid not to oversaturate deep areas, where the signal-to-noise ratio is low. Also, the different angular sectors should have the same contrast, e.g. avoiding a higher gain in the central area.

The acquisitions were performed on 31 regular patients at the CHU in Caen, France, by Dr. Eric Saloux. Figure 6.4 shows a 3D view of one of these acquisitions.

6.4 Registration

Automatic registration is required before the combination step. Registering two (or more) images means finding the geometric transformation T that puts them in correspondence, which can be expressed as the minimisation problem:

$$T = \min_{T} \mu \left(v_1(\mathbf{x}), v_2 \circ T(\mathbf{x}) \right) \tag{6.1}$$

where v_1 and v_2 are the views to be registered and μ is a metric. In the cases were more than two images are to be registered, the standard apical acquisition will be

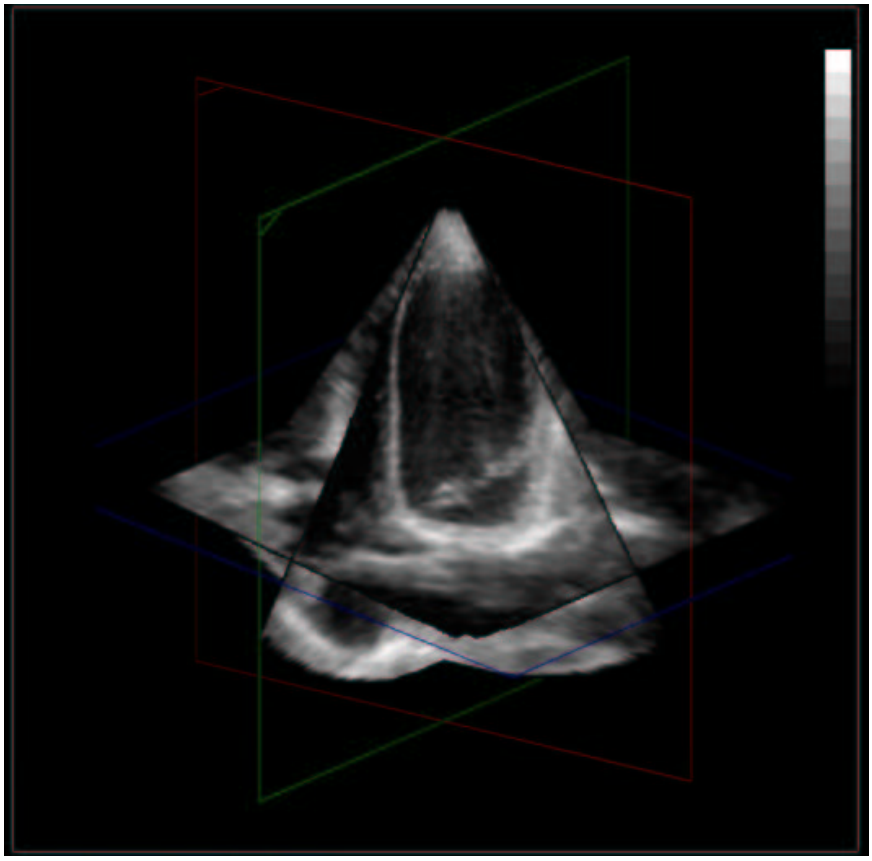


Figure 6.4: Example of a 3D echocardiographic apical acquisition.

used as the target image and all other views will be registered over it. We divide the registration problem into three main components: metric, geometrical transformation and optimization, which we discuss within this section.

6.4.1 Metric

The metric μ in Equation (6.1) has to account for the interesting features to be in correspondence and ignore the rest. The ideal μ function has its peak when images are in correspondence, and gradually lowers its output when images become different. The transition from high to low values depends both on the chosen function and the images, particularly when measuring the correlation between neighbor pixels. For the registration of 3D echocardiographic views, the following points have to be considered:

- Incomplete features. As already mentioned, the walls present in one view may not be visible in the others, due to shadowing or drop-out.
- Imaging artifacts. Close to the probe, important clutter artifacts are present. Also, the SNR gets degraded away from the probe.
- Moving support. The pyramidal shaped domain of the different views makes

the overlapping volume -where μ is measured- very sensitive to the geometrical transformation T.

- Imaging parameters. If the local gains are not properly set (e.g. high gain in the central area), local differences of contrast may mislead the registration.

- Speckle. The presence of speckle makes the direct comparison of the images difficult.

Indeed, registration algorithms need to iterate several times across the transformation space to find the optimal solution. Given the 3D data size, the computational cost of the function is an important parameter.

Another possible approach is to automatically extract segments or landmarks in the image and compute the best transformation between them. The benefits of this technique would be a fast registration algorithm, which focuses on the points of interest, for the same problems that affect iconic registration. Therefore, we focus on iconic methods.

Incomplete features

To avoid over-penalizing for incomplete parts in the different acquisitions, we need a metric that avoids outliers. Many robust estimators (such as maximum likelihood estimators, or M-estimators) have been introduced in the literature (e.g. [126]) for this purpose. Roche *et al.* [128] proposed the generalized correlation ratio with a Geman-McClure M-estimator to register ultrasound and CT images. This ρ-function showed to lead accurate results but poor convergence properties, due to its non-convexity. We propose to use the Huber ρ-function, which gives a good trade off between outlier rejection and convergence. We make the assumption that the intensity levels are the same from the different views, therefore, there is no need to use multimodal similarity metrics such as mutual information or the generalized correlation ratio. We define the metric as:

$$\mu(v_1, v_2 \circ T) = \frac{1}{n(T)} \sum_{\mathbf{x} \in \mathfrak{X}(T)} \rho\left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x})\right) \tag{6.2}$$

where $\mathfrak{X}(T) = \mathfrak{D}(v_1) \cap \mathfrak{D}(v_2 \circ T)$ are the points within the intersection of the domains of the two volumes, $n(T) = \text{card}(\mathfrak{X})$ is the number of pixels in the intersection and the ρ-function is chosen to be the Huber function defined in Equation (2.33). It behaves as a quadratic difference below the threshold α and linearly for higher values. In this way, its derivative is bounded over α, limiting the influence of very different features such as wall and cavity. The α value was chosen to be half way between the maximum and minimum gray level values in $\mathfrak{X}(T)$.

Notice also that the difference between the images is weighted by the number of pixels overlapping. Since the support depends on the transformation, this factor is important to avoid trivial solutions such as a null overlap. Indeed, it normalizes the values of μ so that are comparable between different support sizes.

Masks

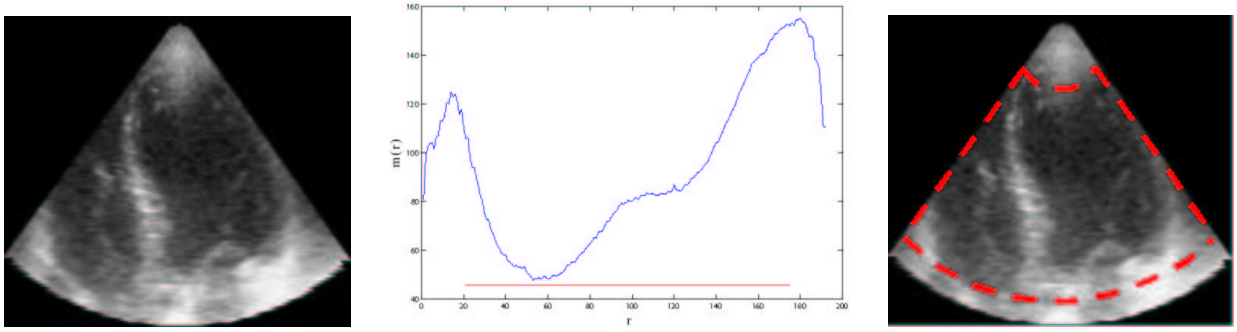
Since the difference between the images is computed with the intensity values, it is important not to consider image artifacts, especially, in the upper part of the images, close to the probe, where some clutter artifacts appear. Also, at the bottom of the images, further from the probe, the gain is usually amplified in order to visualize features, with the inconvenient that signal gets saturated. In order not to take those into account, the actual domain of the images is reduced to the inner part of the acquisition.

To determine the appropriate limits, the radial profile, computed as:

$$m(r) = \frac{1}{\text{card}(\mathcal{S}(r))} \sum_{\mathbf{x} \in \mathcal{S}(r)} v(\mathbf{x}) \tag{6.3}$$

where $\mathcal{S}(r)$ is the subset of points that are at a distance r from the probe position p, within a certain margin ϵ, $\mathcal{S}(r) = \{\mathbf{x} |\ r - \epsilon < d(\mathbf{x}, \mathbf{p}) < r + \epsilon\}$. The two peaks on the radial profile caused by clutter and saturation determine the actual support used for registration.

Figure 6.5.a illustrates the clutter and saturation effects, Figure 6.5.b shows the radial intensity profile $m(r)$ with the two peaks corresponding to clutter and saturation, and Figure 6.5.c shows the actual considered zone.



(a) (b) (c)

Figure 6.5: a) Original image, b) Radial profile of mean value (red line indicates selected domain), c) Image domain used for registration (within dashed line), avoiding clutter and saturation.

6.4.2 Geometrical Transformation

A rigid transformation was chosen according to the following criteria:

- The registration process needs to compensate for the probe motion between the different acquisitions through the different acoustic windows, which can be considered as purely rigid.

- Non-linear deformations might induce "false" motions of the heart walls, which could mislead the diagnosis.

- Different features are present in the different views, therefore non-linear registration would be prone to error in this areas. For instance, if a part of the wall is missing, non-linear registration would enlarge the wall of one view to match the other one.

- A rigid transformation is more stable than affine or non-linear ones.

- Computation of a rigid transformation is much more time efficient due to the limited number of parameters.

- A rigid transformation is enough for the same organ; we can consider that there will not be a significant difference due to a misestimation of the speed of sound because views are close.

A rigid transform can be expressed as a unit quaternion (or versor) $q = a + bi + cj + dk, |q| = 1$ and a translation vector $t = (t_x, t_y, t_z)$. The associated rotation matrix can be expressed as:

$$T = \begin{pmatrix} a^2+b^2-c^2-d^2 & 2bc-2ad & 2ac+2bd & t_x \\ 2ad+2bc & a^2-b^2+c^2-d^2 & 2cd-2ab & t_y \\ 2bd-2ac & 2ab+2cd & a^2-b^2-c^2+d^2 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.4}$$

Using the versor form, the search space is reduced to six parameters, which are directly related to translation and rotation. Compared to the twelve parameters needed for the affine transformation, this guarantees a faster and more robust convergence towards a global minimum.

Data Resampling

In order to evaluate the metric at each step of the transformation, the moving image has to be resampled. We used tri-linear interpolation on the scan converted data. The original data are in polar coordinates, and the scan conversion process consists in resampling them into a Cartesian coordinates. There is a certain degradation in this process, and ideally, each transformation would be performed by a conversion from the original data in order to avoid interpolation artifacts. This would, however, slow down the process.

It is actually the inverse transformation T^{-1}, from the fixed dataset v_0 to the moving image that is computed. This is done to avoid holes in the resampling from the moving image domain to the fixed image domain.

6.4.3 Optimization

We use a gradient descent algorithm to minimize the energy in Equation (6.1). The goal is to find the set of parameters $\gamma = \{a, b, c, d, t_x, t_y, t_z\}$ of the transformation T which minimize the measure μ. Parameters γ are updated in the direction of the steepest gradient as:

$$\begin{aligned} \gamma^0 &= \left\{a^0, b^0, c^0, d^0, t_x^0, t_y^0, t_z^0\right\} \\ \gamma^{t+1} &= \gamma^t - \lambda \frac{\partial}{\partial \gamma} \mu(v_1, v_2 \circ T) \end{aligned} \tag{6.5}$$

Descent direction

The derivation of μ with respect to γ is as follows:

$$\frac{\partial}{\partial \gamma} \mu(v_1, v_2 \circ T) = \frac{\partial}{\partial \gamma} \left(\frac{1}{n(T)} \sum_{\mathbf{x} \in \mathfrak{X}(T)} \rho \left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x}) \right) \right) \tag{6.6a}$$

$$= \frac{\frac{-\partial}{\partial \gamma} n(T)}{n^2(T)} \left(\sum_{\mathbf{x} \in \mathfrak{X}(T)} \rho \left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x}) \right) \right) + \frac{1}{n(T)} \frac{\partial}{\partial \gamma} \left(\sum_{\mathbf{x} \in \mathfrak{X}(T)} \rho \left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x}) \right) \right) \tag{6.6b}$$

Some approximations can be made to simplify this solution. Note that the left term of Equation (6.6b) includes the term $\frac{\partial}{\partial \gamma} n(T)$ in the numerator, which accounts for the variation of the overlapping area in function of the transformation parameters. This value is typically small, especially when compared to the square of the total number of pixels $n(T)^2$ if the images are significantly overlapped. Indeed, it may be difficult to obtain the term $\frac{\partial}{\partial \gamma} n(T)$ for a generic image domain. To compare the importance of the left and right terms of Equation (6.6b), the left term (LT) can be rewritten as:

$$LT = \frac{-1}{n(T)} \frac{\partial n(T)}{\partial \gamma} \mu(v_1, v_2 \circ T) \tag{6.7}$$

And the right term (RT) as:

$$RT = \frac{1}{n(T)} \frac{\partial}{\partial \gamma} \left(n(T) \mu(v_1, v_2 \circ T) \right) = \frac{1}{n(T)} \left(\frac{\partial n(T)}{\partial \gamma} \mu(v_1, v_2 \circ T) + n(T) \frac{\partial}{\partial \gamma} \mu(v_1, v_2 \circ T) \right) \tag{6.8}$$

To evaluate their relative importance, we can develop their ratio as:

$$\left| \frac{RT}{LT} \right| = \left| \frac{n(T)}{\frac{\partial}{\partial \gamma} n(T)} \frac{\frac{\partial}{\partial \gamma} \mu(v_1, v_2 \circ T)}{\mu(v_1, v_2 \circ T)} - 1 \right| \tag{6.9}$$

The left term can be ignored if:

- $n(T)$, the number of pixels in the overlap area, is large.
- $\frac{\partial}{\partial\gamma}n(T)$, the variation of the number of pixels in the overlapping area, is small.
- $\frac{\partial}{\partial\gamma}\mu(v_1, v_2 \circ T)$, the variation of the metric μ, is large.
- $\mu(v_1, v_2 \circ T)$, the metric μ between the two volumes, is small.

The first two conditions are met if the overlap between the two data sets is large. If the third condition is not met, the gradient of μ will be small anyhow. the fourth condition is typically met when the transformation parameters are close to the final solution. On the contrary, in extreme cases where the overlapping area is extremely reduced, the left term could not be ignored. An example would be if the original position of the images overlap for very few pixels; the variation of the weighing term should then be considered.

The right term can be further developed. The characteristic function of the support is defined as:

$$D_v(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \mathfrak{D}(v), \\ 0 & \text{if } \mathbf{x} \notin \mathfrak{D}(v). \end{cases} \tag{6.10}$$

which takes the value 1 inside $\mathfrak{D}(v)$, the support of volume v, and 0 outside. The overlapping support of the M images to register is defined as:

$$D(\mathbf{x}) = \prod_{i=1..M} D_{v_i \circ T_i}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \cap_{i=1..M}\mathfrak{D}(v_i \circ T_i), \\ 0 & \text{if } \mathbf{x} \notin \cap_{i=1..M}\mathfrak{D}(v_i \circ T_i). \end{cases} \tag{6.11}$$

We can rewrite and develop the right term of Equation (6.6b) as:

$$\begin{aligned} RT &= \frac{1}{n(T)}\frac{\partial}{\partial\gamma}\sum_{\mathbf{x}\in\mathfrak{X}(T)}\rho\left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x})\right) \\ &= \frac{1}{n(T)}\frac{\partial}{\partial\gamma}\sum_{\mathbf{x}\in\mathfrak{U}}\rho\left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x})\right) D(\mathbf{x}) \\ &= \frac{1}{n(T)}\sum_{\mathbf{x}\in\mathfrak{U}}\frac{\partial}{\partial\gamma}\left[\rho\left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x})\right) D(\mathbf{x})\right] \\ &= \frac{1}{n(T)}\sum_{\mathbf{x}\in\mathfrak{U}}\frac{\partial}{\partial\gamma}\left[\rho\left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x})\right)\right] D(\mathbf{x}) + \rho\left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x})\right)\frac{\partial}{\partial\gamma}D(\mathbf{x}) \\ &= \frac{1}{n(T)}\sum_{\mathbf{x}\in\mathfrak{X}(T)}\frac{\partial}{\partial\gamma}\rho\left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x})\right) + \sum_{\mathbf{x}\in\mathfrak{U}}\rho\left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x})\right)\frac{\partial}{\partial\gamma}D(\mathbf{x}) \end{aligned} \tag{6.12}$$

where $\mathfrak{U}$ is the universe of all pixels, and does not depend on T. Assuming that images are registered by pairs, that is, that there is only one moving image, the term $\frac{\partial}{\partial\gamma}D(\mathbf{x})$

can be further developed as:

$$\frac{\partial}{\partial\gamma}D(\mathbf{x}) = \frac{\partial}{\partial\gamma}D_{v_1}(\mathbf{x})D_{v_2}(T(\mathbf{x})) = D_{v_1}(\mathbf{x})\nabla D_{v_2}(T(\mathbf{x}))\frac{\partial}{\partial\gamma}T(\mathbf{x}) \tag{6.13}$$

The gradient $\nabla D_{v_2}(T(\mathbf{x}))$ only takes different values form zero at the boundaries of $\mathfrak{D}(v_2)$. Therefore, the right term of Equation (6.12) is only evaluated at the points $\mathbf{x} \in \mathfrak{D}(v_1) \cap \mathfrak{F}(\mathfrak{D}(v_2 \circ T))$ where $\mathfrak{F}(\bullet)$ represents the points on the boundary. We can now rewrite RT as:

$$RT = \frac{1}{n(T)} \sum_{\mathbf{x}\in\mathfrak{X}(T)} \frac{\partial}{\partial\gamma}\rho\left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x})\right) + \sum_{\mathbf{x}\in\mathfrak{D}(v_1)\cap\mathfrak{F}(\mathfrak{D}(v_2\circ T))} \rho\left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x})\right) \nabla D_{v_2}(T(\mathbf{x}))\frac{\partial}{\partial\gamma}T(\mathbf{x}) \tag{6.14}$$

Again, if we consider that $\text{card}(\mathfrak{D}(v_1) \cap \mathfrak{F}(\mathfrak{D}(v_2 \circ T)) \ll n(T)$, that is, that the number of points within the border of the support of the moving image and within the support of points of the fixed image is much smaller than the total overlapping points, we can neglect this second summation term. Therefore, the direction of the gradient is simplified to:

$$\frac{\partial}{\partial\gamma}\mu(v_1, v_2 \circ T) \approx \frac{1}{n(T)} \sum_{\mathbf{x}\in\mathfrak{X}(T)} \rho'\left(v_1(\mathbf{x}) - v_2 \circ T(\mathbf{x})\right) \nabla(v_2 \circ T(\mathbf{x})) \cdot J(T) \tag{6.15}$$

where $\rho'(\bullet)$ is the derivative of the $\rho(\bullet)$ function, $\nabla(\bullet)$ the gradient vector and $J(\bullet)$ the Jacobian matrix. In fact, with these approximations we are not considering the effect that the overlapping area changes in function of the transformation, which is reasonable if the overlapping area is large enough and parameters of the transformation changes are small.

Initialization

Gradient descent algorithms converge to local minima. Since the metric μ cannot be guaranteed to be convex with respect to the transformation parameters, a proper initialization is needed to reach the global minimum corresponding to the solution of Equation (6.1). Indeed, initialization will ensure a sufficient overlap between the different volumes which permits to use the aforementioned simplifications.

To obtain a robust initialization, it is necessary to know some correspondance between the different views. However, it is actually quite difficult to locate precisely, in a repeatable way, the same anatomical positions in the different volumes. This is due to several reasons, including the presence of noise and the similarity of the different cuts due to the rotational symmetry of the internal heart structures. Therefore, we chose another way to initialize, by making use of an automatic border detection method in each view. Five points were manually selected on each view: two in the mitral annulus in the four-chamber short-axis view, two in the mitral annulus in the two-chamber

view and one on the apex. From this points, a segmentation of the left ventricule is generated. This segmentation does not need to be perfect; it serves as starting point to perform a precise registration. The segmentation was performed with the commercial software QLab (Philips, Best, The Netherlands), generating a mesh of about 600 points. The optimal rigid transformation between two meshes was then found by minimizing the summed squared distance between all point coordinates. We denote this procedure as *mesh-based* initialization.

The correspondence between the points is known. Denoting by $\mathbf{R}_{ini}$ the initial rotational matrix, by $\mathbf{T}_{ini}$ the translational part, $\mathbf{p} = \{p_j\}$ the reference set of points, $\mathbf{p}' = \{p'_j\}$ the set of points of the moving image and N the number of points in the mesh, we want to minimize the following expression:

$$(\mathbf{R}_{ini}, \mathbf{T}_{ini}) = \min_{\mathbf{R}_{ini}, \mathbf{T}_{ini}} \sum_{j=1}^{N} \left\| p_j - (\mathbf{R}_{ini} p'_j + \mathbf{T}_{ini}) \right\|^2 \tag{6.16}$$

By defining $q \equiv p - p_c = p - \frac{1}{N}\sum_j p_j$ and $q' \equiv p' - p'_c = p' - \frac{1}{N}\sum_j p'_j$ as the mesh points subtracting their center of masses (p_c and p'_c respectively), we can develop Equation (6.16) as:

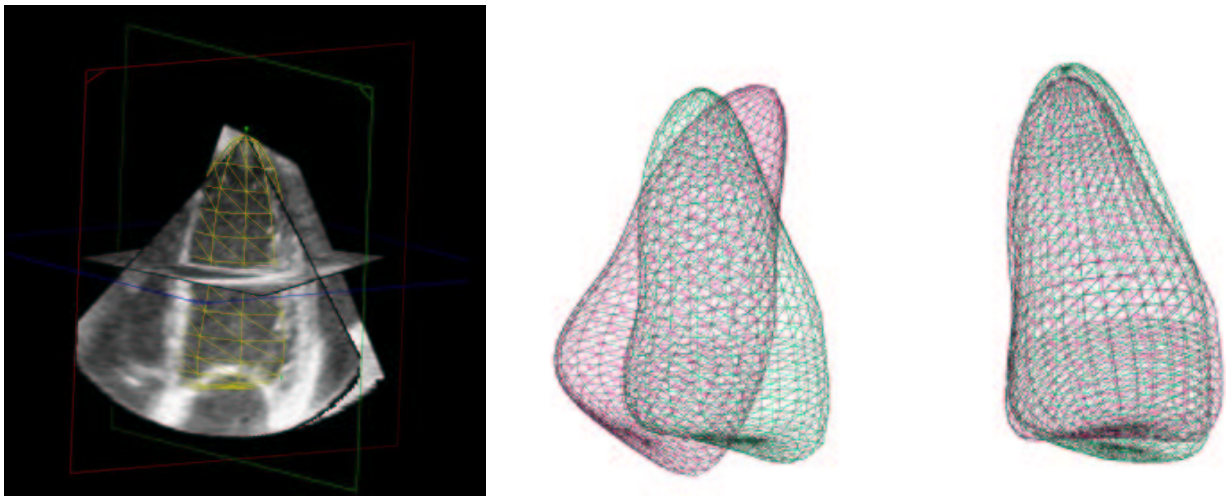
$$\begin{aligned} \mathbf{R}_{ini} &= \min_{\mathbf{R}_{ini}} \sum_{j=1}^{N} \left\| q_j - \mathbf{R}_{ini} q'_j \right\|^2 \\ \mathbf{T}_{ini} &= p_c - \mathbf{R}_{ini} p'_c \end{aligned} \tag{6.17}$$

where the rotation matrix $\mathbf{R}_{ini}$ can be obtained through singular value decomposition (SVD) or using the quaternion approach.

Figure 6.7 shows the points where the user is asked to click. Both views should be aligned so that the points are approximately in correspondence. Figure 6.6.a shows an example of a segmented ventricle on the intensity image, along with two segmented ventricles at their original locations (Figure 6.6.b) and after their registration (Figure 6.6.c). Other approaches, such as iterative closest point (ICP) [24] could be used with the benefit that there is no need to know in advance the correspondence between the points on the meshes. The egg-shape of the ventricle may induce unwanted rotations, which we want to explicitly avoid by making use of the one-to-one correspondence between the points of the mesh. The transformation obtained from the meshes is used as initialization of the transformation between the two acquisitions, as starting point for the automatic gradient descent.

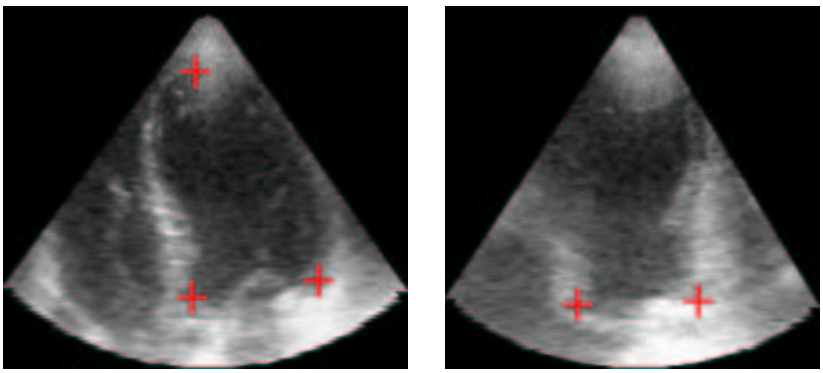
Multi-resolution Approach

In order to avoid local minima and to speed up computation time, a multiresolution scheme is used. It consists in blurring -to avoid aliasing- and downsampling the



(a) (b) (c)

Figure 6.6: a) Example of segmented left ventricle, b) meshes at their original location, c) after initial registration.



(a) (b)

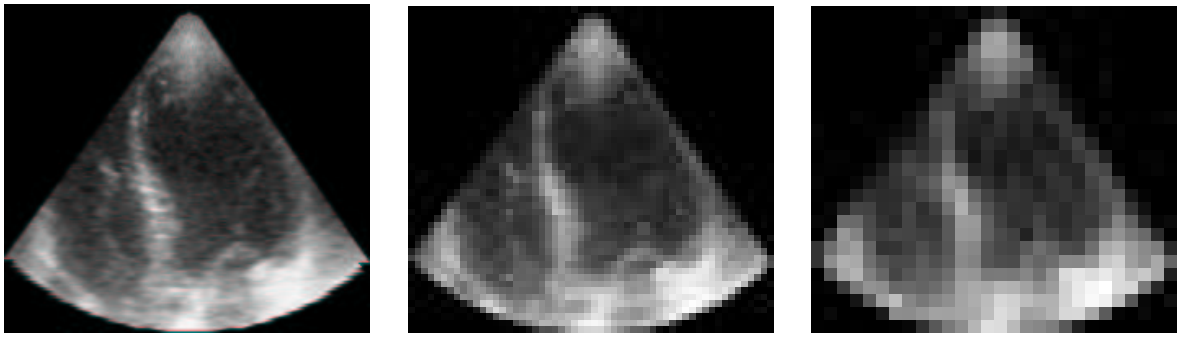
Figure 6.7: Points to click by the user. a) four chamber short-axis view, b) Two-chamber view.

different views at different levels l as:

$$\begin{aligned} \sigma_l &= \left(\frac{l}{2}\right)^2 \\ v^l(\mathbf{x}) &= v(l\mathbf{x}) * \frac{1}{\sigma_l\sqrt{2\pi}} e^{-(l\mathbf{x})^2/(2\sigma_l^2)} \end{aligned} \tag{6.18}$$

where the convolution process takes place before the downsampling. The number of voxels of the volumes is reduced by a factor of l^3 assuming that the initial scale is $l = 1$. Typically the levels are taken as powers of 2, $l = 2^s$. Figure 6.8 shows a cut plane at different resolution levels. The registration process starts at the coarser level, and uses the found transformation as starting point for the next level. Figure 6.9 shows

the evolution of the metric μ along the iterations.



(a) (b) (c)

Figure 6.8: Four chamber view at different multi-resolution levels. a) $l = 1$ (original resolution), b) $l = 4$, c) $l = 8$.

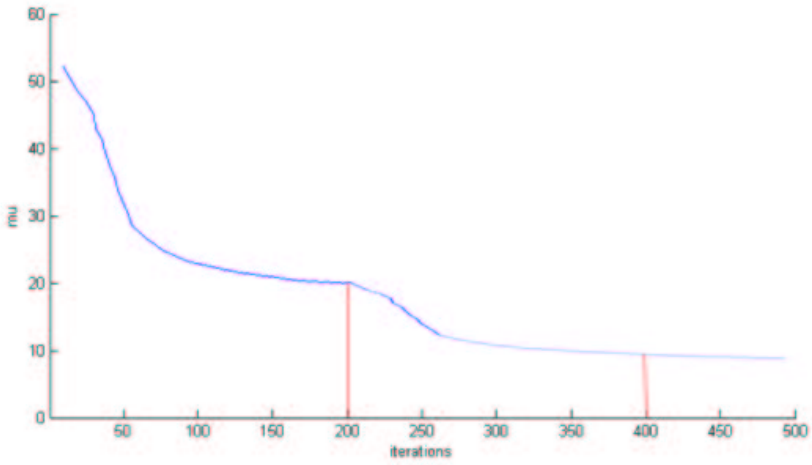


Figure 6.9: Evolution of the μ metric along the iterations. Red lines indicate a change in scale.

Temporal Alignment

Registration of different views acquired through different cardiac cycles relies on the hypothesis that the heart movements are consistently periodic. This is a strong hypothesis, which does not hold in all the cases; but it is reasonable. First of all, with the current 3D echocardiographic technology, each "full volume" acquisition is obtained through four consecutive cardiac cycles; one fourth of the whole sector being acquired during each cardiac cycle. If the heart is not stable enough through these four cycles, artifacts will appear and the acquisition will be rejected. During the exam, patients are asked not to breath, in order to minimize movement artifacts. Acquisitions used

in our experiments showed a consistent movement within acquisition time. A second consideration that needs to be taken into account is that pathologies may potentially show motion irregularities through the different cycles (e.g. arrhythmias), can be triggered by pacemaker and/or relaxed with beta blockers. Therefore, even if the non-periodicity of heart motion is a potential issue, it is controllable within bounded limits. The different acquisitions are triggered by the electrocardiogram, which ensures the temporal synchronization of at least the initial frame. Shekhar *et al.* [134], in the context of stress echocardiography, propose a piecewise linear temporal scaling to align pre-stress and post-stress acquisitions, with a linear scaling of the systole cycle and another linear scaling of the diastole cycle. The end diastole is determined by the ECG, and the end-systole frame was determined with the user intervention. In our study, we supposed that the cardiac cycle was constant.

The registration process was performed for the first frame of the sequence (end-diastole, triggered by the ECG), and the same T was then used for the whole cardiac cycle. This showed to be a reasonable approximation. Further refinnement could be obtained by performing the registration process at each frame.

6.4.4 Registration Validation

To validate the accuracy of the registration method, we have compared the obtained transformation with (i) a known transformation on phantom measurements, (ii) visually on regular *in vivo* exams and (iii) a transformation measured with an optical position sensor on human patients.

Phantom data

The ATS 539 tissue-mimicking phantom (see appendix B) was scanned with a mechanical arm each 2 cm through the scanning surface #1. The precision of the mechanical arm is on the order of 1 μm, significantly below the resolution of the images. Measurements with different angulations in the lateral plane were also performed, and aligned manually afterward. To evaluate the precision achieved by the registration algorithm, the final transformation was compared to the measured one. Transformations Δt were computed between 10 consecutive acquisitions that were laterally translated with a 2 cm step, and $\Delta\alpha$ between 3 acquisitions for angular displacement of 15 degree. Table 6.2 reports error measurements on these phantom data, for lateral displacements Δt and angular displacements $\Delta\alpha$. No mesh-based initialization was performed on phantom data. Figure 6.10 shows the phantom volumes before and after registration.

Δt	$0.68mm \pm 0.18$
$\Delta\alpha_x$	$2.29° \pm 0.82$

Table 6.2: Registration errors on tissue-mimic phantom data.

It must be considered that the phantom has large tubular structures, which are dif-

ferent from the heart anatomy. These values are an upper bound of the achievable performance on *in vivo* data.

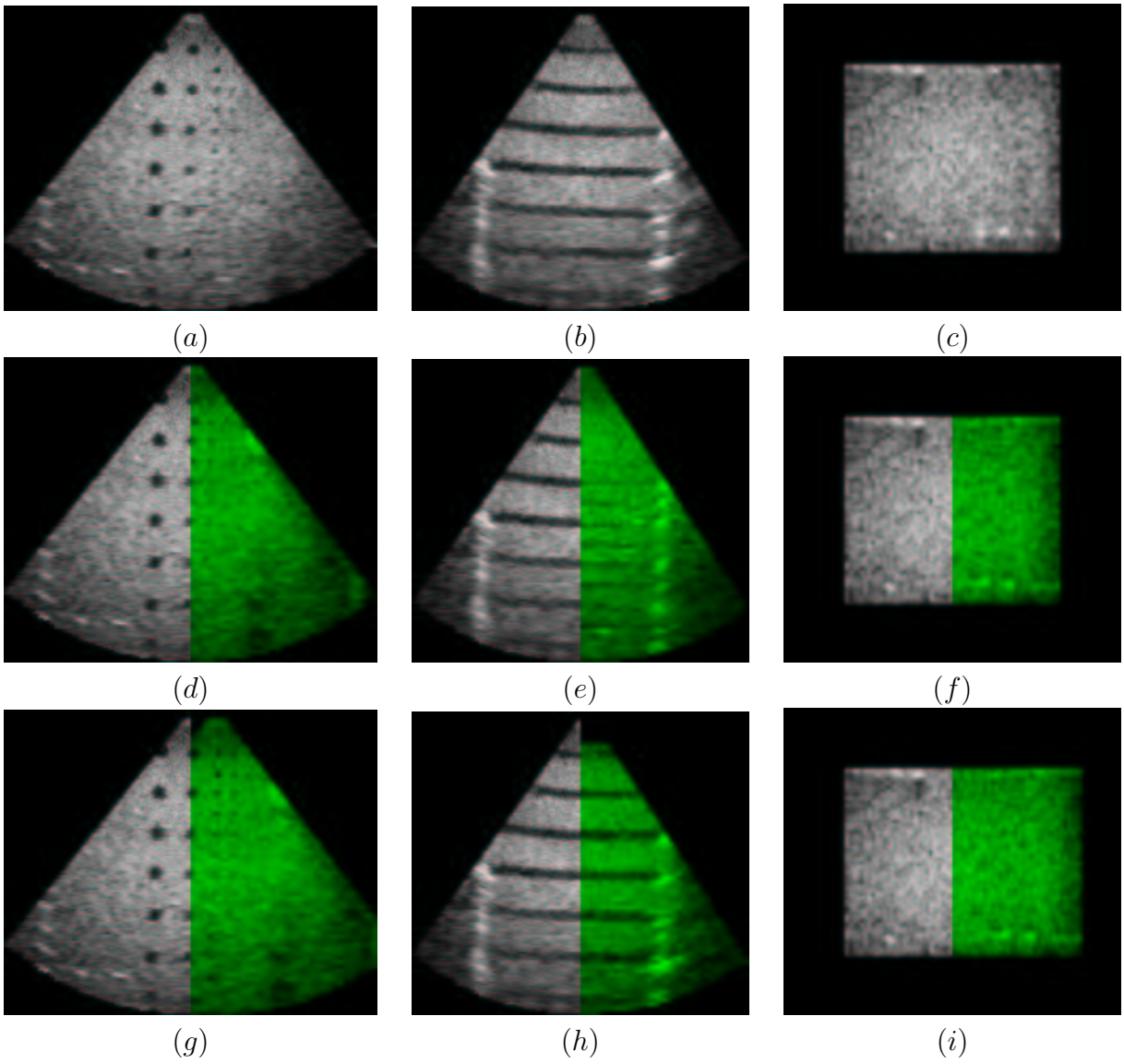


Figure 6.10: Phantom registration of translated acquisition. Reference volume: a) B-scan, b) E-scan , c) C-scan. Initial phantom position (right - green) and reference volume (left - gray): d) B-scan, e) E-scan , f) C-scan. Automatically registered volume (right - green) and reference volume (left - gray): g) B-scan, h) E-scan , i) C-scan.

Visual evaluation

Figure 6.11 illustrates the registration results for one *in vivo* data set. It can be seen how the septum wall, misaligned in the original volumes, is better aligned in the initialized volumes, but with a considerable rotation difference in the short axis view.

Similarly, other areas not shown in these figures are misaligned [1]. When the volumes are automatically registered, most deviations are corrected. The same transformation is then applied to the whole cardiac cycle.

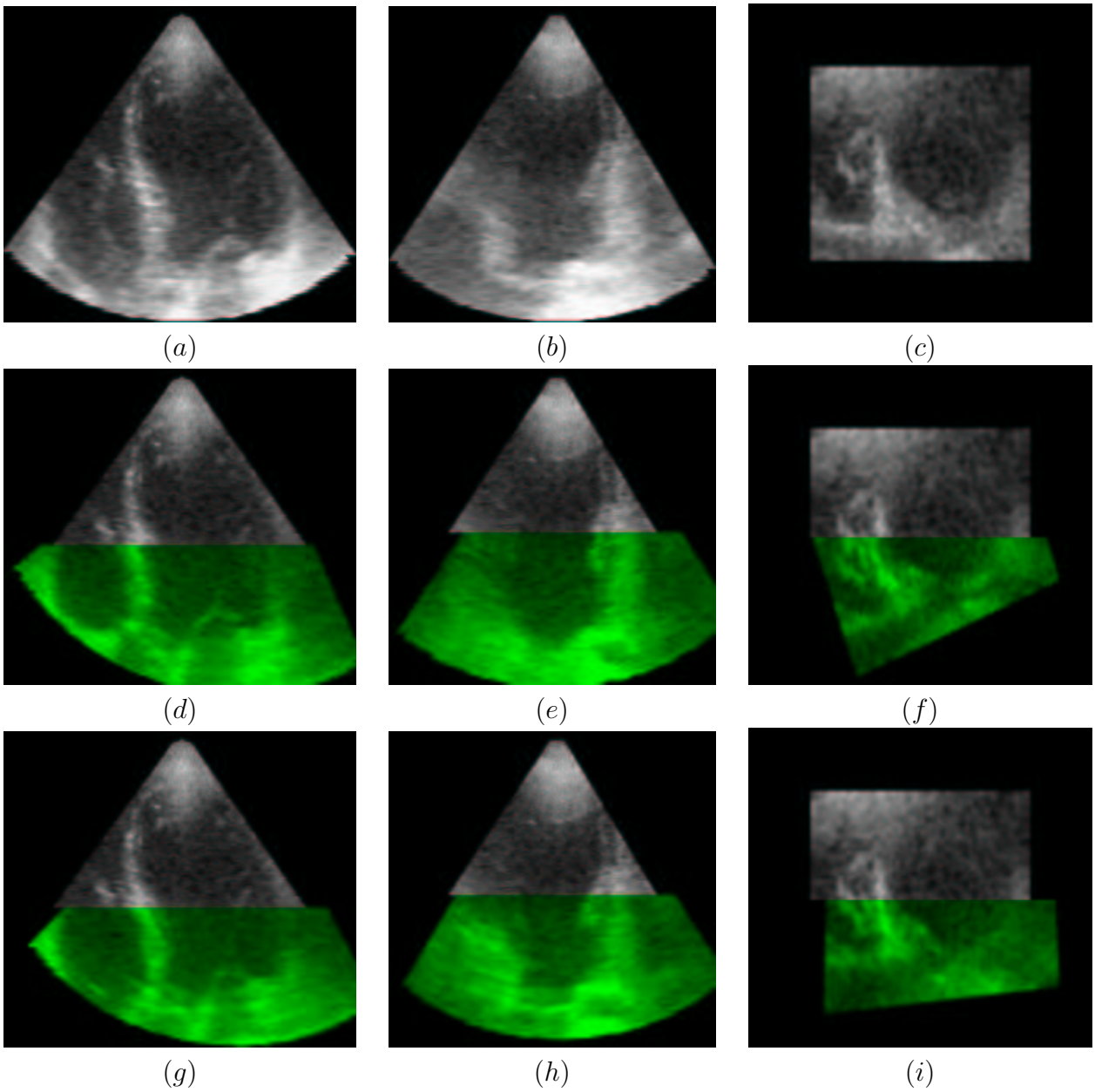


Figure 6.11: Registration initialization example. Reference volume: a) B-scan, b) E-scan , c) C-scan. Manually initialized volume (bottom - green) and reference volume (top - gray): d) B-scan, e) E-scan , f) C-scan. Automatically registered volume (bottom - green) and reference volume (top - gray): g) B-scan, h) E-scan , i) C-scan.

[1] Consider the difficulty of displaying three dimensional volumes on paper.

Optical Tracker

Acquisitions on regular human patients were obtained with two different probes mounted with an optical marker to track their position. The relative position of the probes was recorded by an external Polaris(R) optical sensor, with an accuracy precision 0.35 mm RMS. Acquisitions were made from the apical windows, separated approximately by two centimeters. The data for these experiments was obtained on two different patients at the Massachusetts General Hospital of Boston, MA, USA.

Table 6.3 summarizes the differences between the transformation from the sensor and those obtained with the registration algorithm, in terms of translation (Δt) and rotation in the lateral ($\Delta\alpha_x$), axial ($\Delta\alpha_y$) and elevational ($\Delta\alpha_z$) dimensions. As observed, the accuracy is lower than in the phantom experimentation. The worse accuracy was obtained in the axial rotation axis, probably due to the auto-similarity of the heart cavities within this axis.

Δt	$3.18mm \pm 0.94$
$\Delta\alpha_x$	$3.38° \pm 0.57$
$\Delta\alpha_y$	$10.94° \pm 1.12$
$\Delta\alpha_z$	$8.16° \pm 1.85$

Table 6.3: Versor difference between the transformation obtained with an optical tracker and the registration algorithm. Translation (Δt) and rotation in the lateral ($\Delta\alpha_x$), axial ($\Delta\alpha_y$) and elevational ($\Delta\alpha_z$) dimensions.

6.5 Restoration Results

In this Section we describe the results on dataset acquired on patients using the techniques commented so far. Particularities on these techniques will be described in the first subsection. To facilitate the description, Figure 6.12 shows the used nomenclature for the different heart segments.

6.5.1 Multiview Restoration Methods

In this Section we summarize the restoration techniques used to combine the different cardiac views. No modifications were introduced in the generalized averages, maximum frequency and wavelet fusion techniques introduced in Chapter 4. On the other hand, in this section we describe the modifications that apply to weighted averaging and multiview deconvolution due to the different needs of the application. Table 6.4 summarizes the presented techniques.

Maximum

Since the main objective of this application is to keep all heart walls, which appear as bright areas on the images, we found it convenient to evaluate the performance of the

Method	Equation	Section	Parameters
Maximum (mx)	$v_{max}(\mathbf{x}) = \max\left(v_1(\mathbf{x}), \ldots, v_M(\mathbf{x})\right)$	6.5.1	-
Weighted Average (wa)	$v_{wa}(\mathbf{x}) = \sum_{i=1}^{M} \alpha_i(\mathbf{x})$	4.2.2	α_i, σ_n, σ_s
Generalized Average (ga)	$v_{ga}(\mathbf{x}) = \left(\frac{1}{M}\sum_{i=1}^{M} v_i(\mathbf{x})^{\beta(\mathbf{x})}\right)^{1/\beta(\mathbf{x})}$	4.2.3	$\beta(\mathbf{x})$, σ_n, σ_s, γ
Maximum Frequency (mf)	$V_{mf}(\mathbf{f}) = max_i\left\{V_i(\mathbf{f})\right\}$	4.3.1	-
Wavelet Fusion (wf)	$\tilde{w}^l(\mathbf{y}) = \max_i\left\{\alpha_i(\mathbf{y}) w_i^l(\mathbf{y})\right\}$	4.3.2	α_i, σ_n, k
Multiview Deconvolution (md)	$v_{md} = \arg\min_{v_{md}}\left[(1-\lambda)\sum_{i=1}^{M}\alpha_i \left\|v_i - h_i * v_{md}\right\|^2 + \lambda\Psi(v_{md})\right]$	3	$\mathbf{h}$, σ_n, σ_s, λ

Table 6.4: Restoration Methods for cardiac imaging.

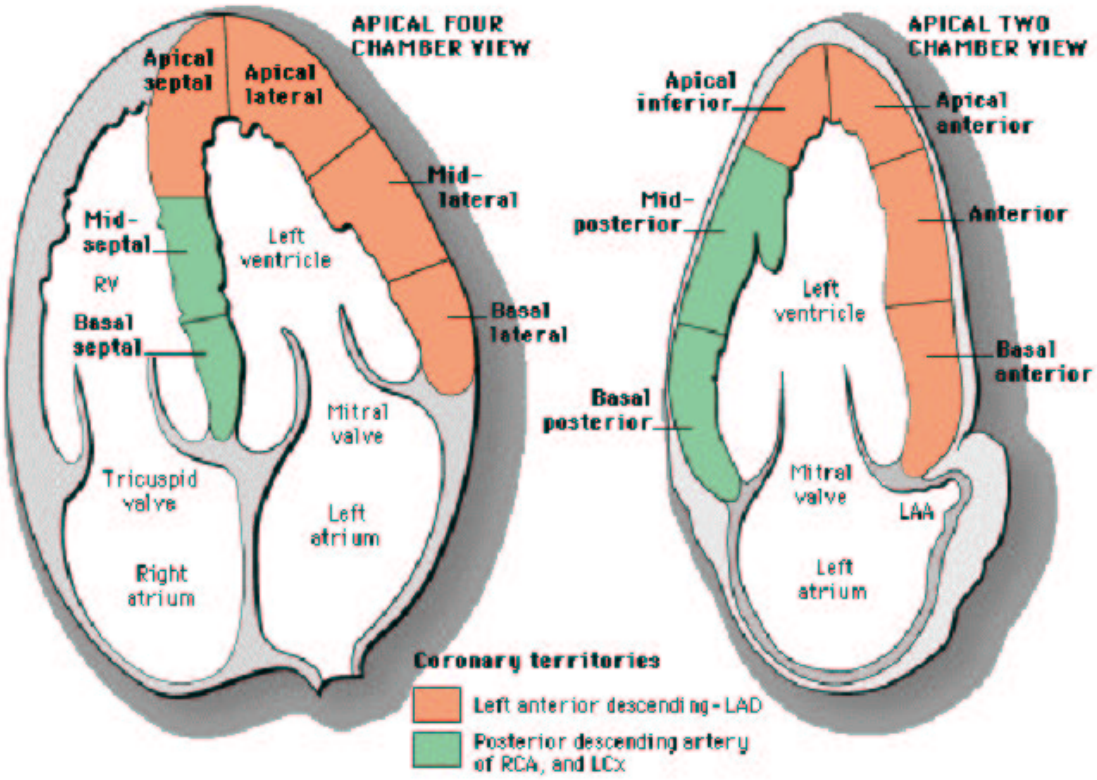


Figure 6.12: Heart Segment description[8].

maximum operator. The maximum volume is defined as the maximum value at each pixel:

$$v_{max}(\mathbf{x}) = \max\left(v_1(\mathbf{x}), \ldots, v_M(\mathbf{x})\right) \tag{6.19}$$

Weighted Averaging

A more adapted saliency measure can be used for detecting heart walls than was needed for the breast application. Since walls are typically much larger than speckle patterning, we relaxed the saliency measure to simply detect the image areas with high intensity, using directly its intensity value, as:

$$\alpha_i^{'}(\mathbf{x}) = sal_i(\mathbf{x}) = \bar{v_i} \tag{6.20}$$

where $\bar{\bullet}$ represents a local mean. The final volume is computed as:

$$\alpha_i = \frac{\alpha_i^{'}}{\sum_i \alpha_i^{'}} \tag{6.21a}$$

$$v_{wa}(\mathbf{x}) = \sum_{i=1}^{M} \alpha_i(\mathbf{x}) v_i(\mathbf{x}) \tag{6.21b}$$

Multiview deconvolution

Given the size of the features of interest, its presence or absence may be due to shadowing rather than to the spatial sensitivity of the acquisition system. To compensate for this effect, we introduce the saliency measure used in the weighted averaging into the multichannel deconvolution framework, as:

$$v_{md} = \arg\min_{v_{md}} \left[(1-\lambda) \sum_{i=1}^{M} \alpha_i \left\| v_i - h_i * v_{md} \right\|^2 + \lambda \Psi(v_{md}) \right] \tag{6.22}$$

With respect to the estimation of the point spread function, some further constraints have been introduced. The reason is that there is not much diversity between the different views, which in some cases are almost parallel, and indeed, the obtained images are very noisy. The different h_i are assumed to be Gaussian, with variance linearly depending on the distance to the probe r:

$$h_i = G(\mathbf{0}, a \cdot r(\mathbf{x})) \tag{6.23}$$

In practice, the volume was separated in slices of 1 cm of thickness, where the point spread function h was supposed to be constant. The parameter a was estimated using the cross-channel relationship as:

$$a = \arg\min \left\| h_1 * v_2 - h_2 * v_1 \right\|^2 \tag{6.24}$$

6.5.2 Multiple Apical Acquisitions

From the acquired data, exams on eight different patients were chosen as representative set of this application, with average echographic quality and complementarity between the views. The goals, as mentioned, are to increase the border visibility, the signal-to-noise ratio and the spatial resolution.

Quantification

To quantitatively evaluate the restoration processes, we defined the myocardium contrast increase as:

$$\Delta C = \left(\left| \frac{\bar{M}_f - \bar{B}_f}{\bar{M}_1 - \bar{B}_1} \right| - 1 \right) \cdot 100 \tag{6.25}$$

where $\bar{M}$ is the mean value in a manually defined window on the lateral, anterior and posterior myocardium walls, and $\bar{B}$ the mean value in a manually defined window in the LV cavity. The improvement of signal-to-noise ratio is defined as:

$$\Delta SNR = 20 \log \frac{\sigma_1}{\sigma_f} \tag{6.26}$$

where σ is the noise standard deviation in manually defined windows in the LV cavity, which is supposed to have a constant value. Table 6.5 shows the value of these parameters with the proposed fusion techniques.

	mx	wa	ga	mf	wf	md
ΔC	+39.0%	+37.3%	+45.3%	+45.6%	+47.0%	+49.7%
ΔSNR	+3.87	+4.91	+4.74	-1.11	+4.67	+4.61

Table 6.5: Image quality improvement parameters for the different fusion techniques, maximum (mx), weighted average (wa), generalized average (ga), maximum frequency (mf), wavelet fusion (wf) and multiview deconvolution (md).

The field of view, measured as:

$$FOV = \left(\frac{n_f}{n_1} - 1\right) \cdot 100 \tag{6.27}$$

was improved by +18% on average for two acquisitions. This value depends uniquely on the position of the probe, thus, it is independent of the fusion technique.

Visual evaluation

Figure 6.13 shows two apical windows, a standard (named v_0), and one displaced by about 2 cm to the lateral side, towards the patient's left arm (named v_1). The mid and apical lateral segments fall nearly outside the field of view of the standard acquisition. The second acquisition exhibits better contrast in those areas. However, the apex falls outside the field of view of this second acquisition. As appreciated in the short-axis view (Figure 6.13 e and f), the mid septal anterior segment shows a poor contrast in v_0 and better in v_1. The endocardium is difficult to distinguish in all views, although that is somewhat clearer in v_0 in the four-chamber and two-chamber views, next to the lateral wall.

Figures 6.14, 6.15 and 6.16 show the result of fusion on three different slices (four-chamber, two-chamber and short axis) for the following techniques: maximum, weighted average, generalized average, maximum frequency, wavelet fusion and multiview deconvolution. In general lines, maximum, weighted averaging and generalized averaging behave quite similarly, although the second shows smoothed patterns in the blood pool. Since the views are very different in terms of spatial resolution, it is not possible to improve it by means of generalized averaging. Maximum frequency shows lots of artifacts, although walls show a strong contrast. Wavelet fusion and multiview deconvolution also show a very similar behavior in all planes, although the contrast of features is slightly higher for the multiview deconvolution. However, the noise patterns appear more accentuated in these two techniques, as well as wall borders, which appear smoother in the averaging techniques. From now on, we will focus our description on the maximum, which is the simplest technique, and the multichannel deconvolution techniques, which performs best.

Figure 6.17 shows a zoom on an arbitrary oblique plane, showing the improvement in the visibility of the septum and the lateral wall. Also notice how the maximum technique tends to enlarge the walls, and how the multichannel deconvolution technique stays closer to the original size. Figure 6.18 and 6.19 show similar results on a zoom

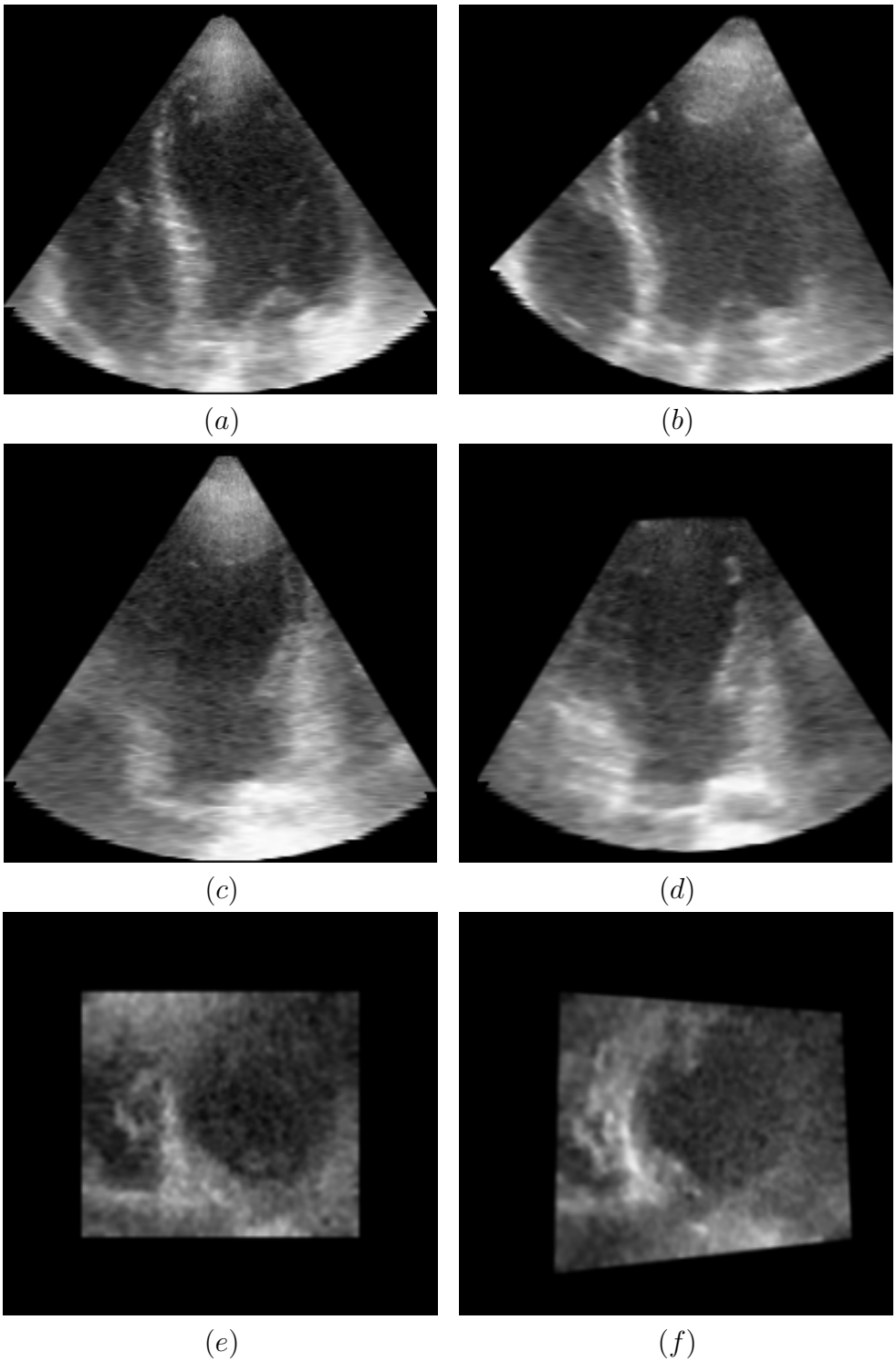


Figure 6.13: Original apical views. Left, standard apical v_0 from top to bottom: a) four-chamber view, c) two chamber view, e) short-axis view. Right, exterior apical v_1 from top to bottom: b) four-chamber view, d) two chamber view, f) short-axis view.

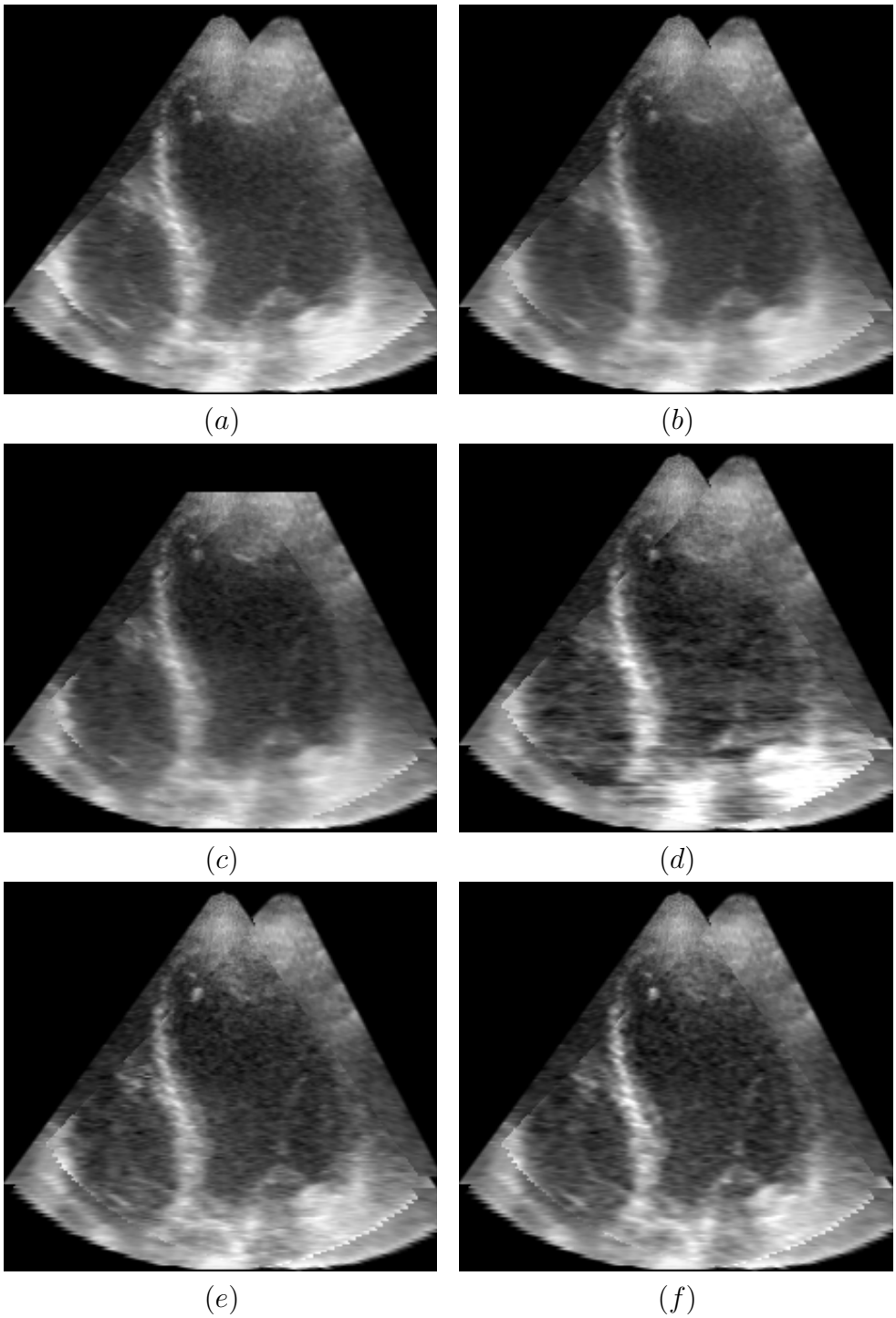


Figure 6.14: Four-chamber apical view of combined views for *in vivo* data. a) Maximum, b) weighted average, c) generalized average, d) maximum frequency, e) wavelet fusion, f) multiview deconvolution.

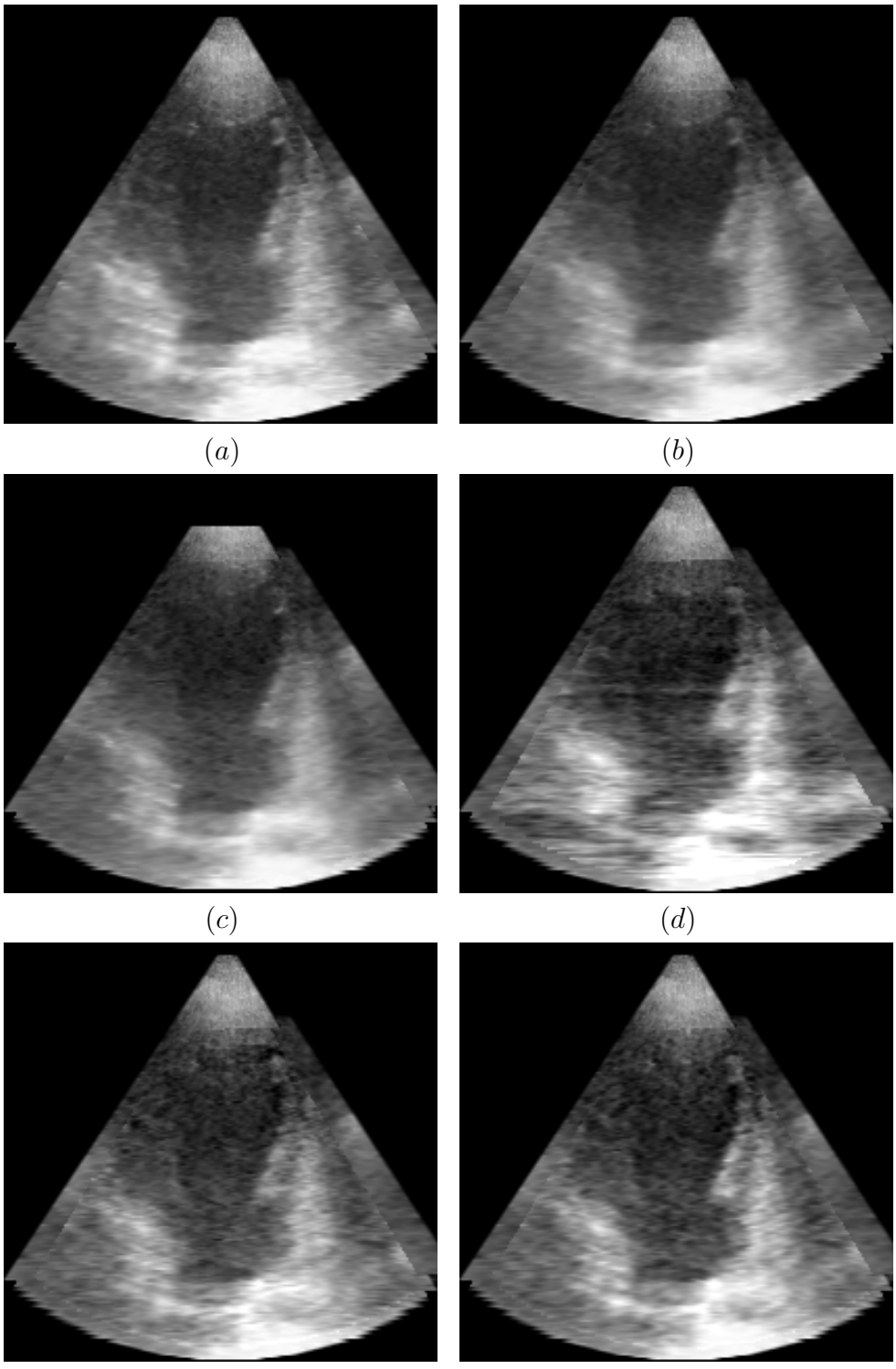


Figure 6.15: Two-chamber apical view of combined views for *in vivo* data. a) Maximum, b) weighted average, c) generalized average, d) maximum frequency, e) wavelet fusion, f) multiview deconvolution.

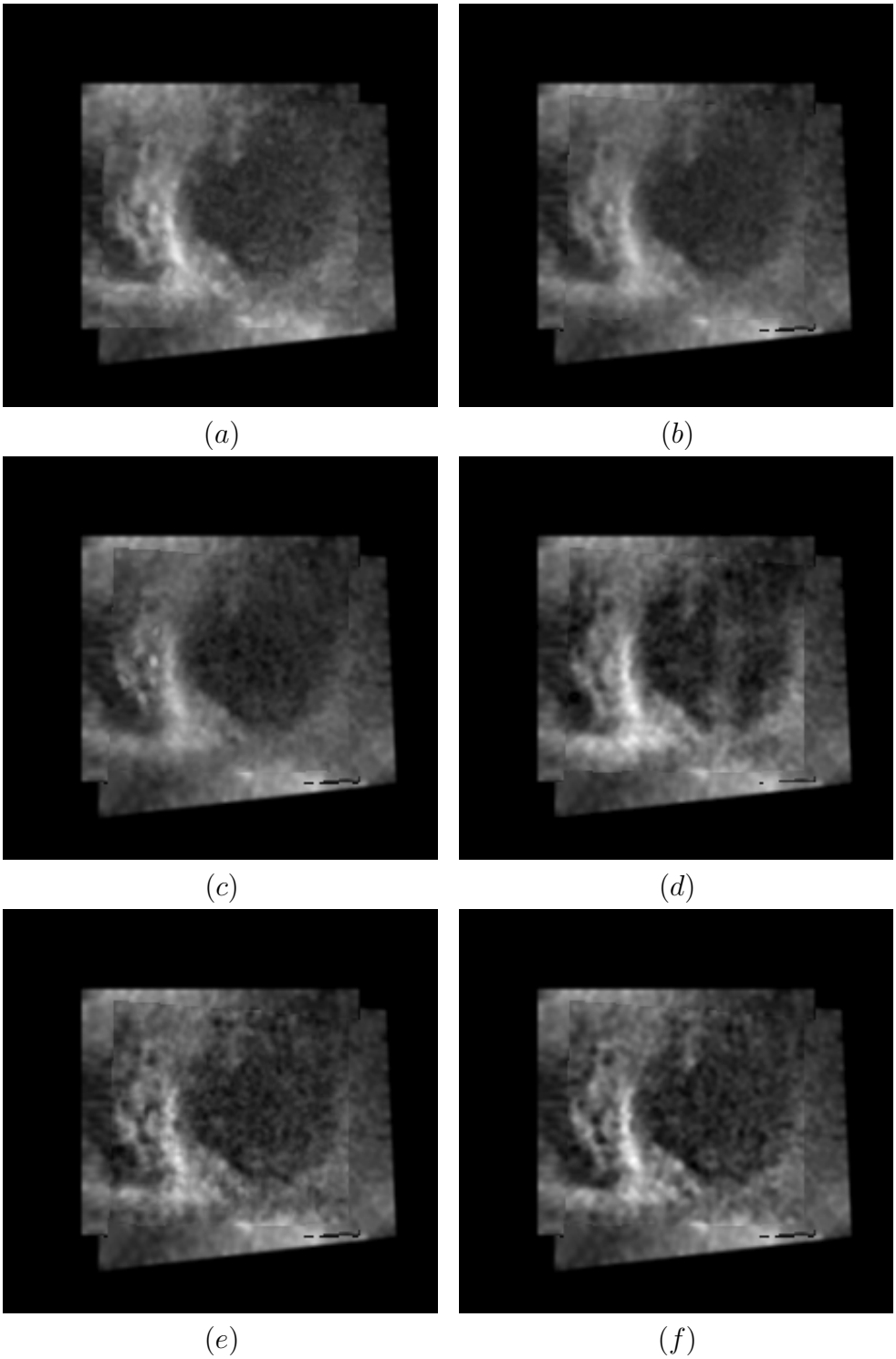
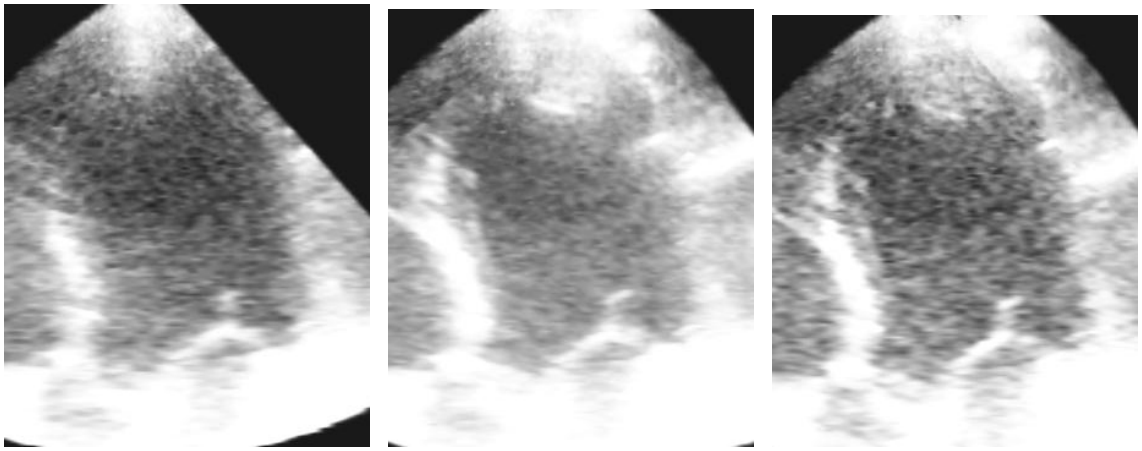


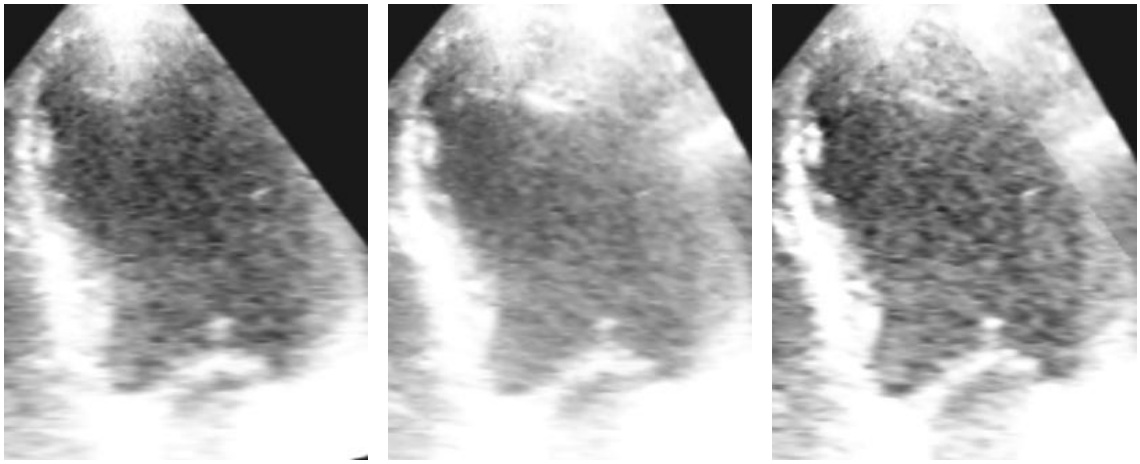
Figure 6.16: Short-axis apical view of combined views for *in vivo* data: a) maximum, b) weighted average, c) generalized average, d) maximum frequency, e) wavelet fusion, f) multiview deconvolution.

on the left ventricle and a short axis view. In the latter, notice how the wall contrast of the right ventricle has been improved. Again, note how the maximum technique enlarges the walls in comparison to others.



(a) (b) (c)

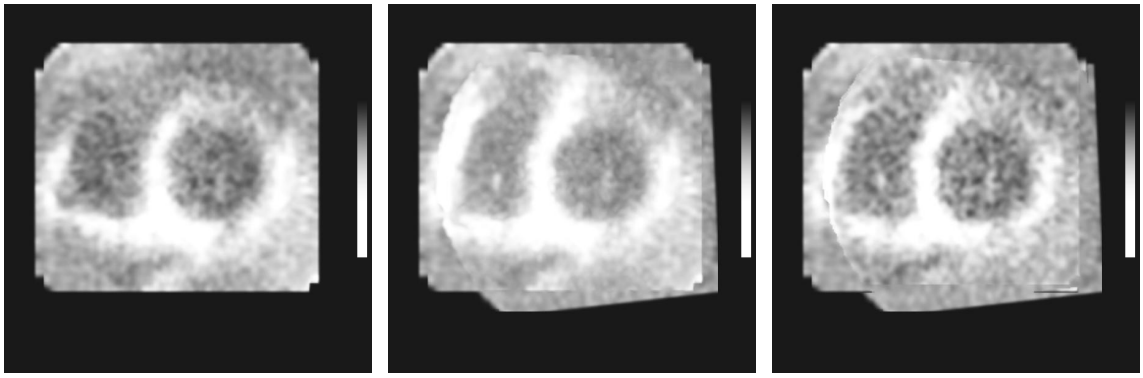
Figure 6.17: Zoom on the left ventricle on an arbitrary oblique plane, with particular interest on the septum and the lateral walls: a) v_0, b) maximum, c) multiview deconvolution.



(a) (b) (c)

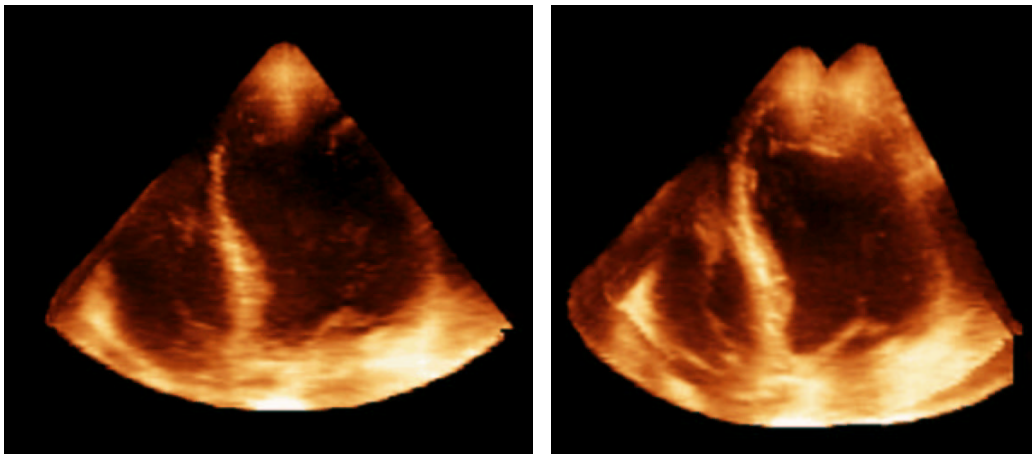
Figure 6.18: Zoom on the left ventricle on an arbitrary oblique plane, with particular interest on the delimitation of the left ventricle. a) v_0, b) maximum, c) multiview deconvolution.

Figure 6.20 shows a three-dimensional reconstruction of a standard apical acquisition and combined views with multiview deconvolution technique. Notice how myocardium walls are more visible on combined views, especially at the apex and the lateral wall.



(a) (b) (c)

Figure 6.19: Zoom on a mid short axis view with particular interest on the right ventricle. a) v_0, b) maximum, c) multiview deconvolution.



(a) (b)

Figure 6.20: Three-dimensional visualization of cardiac data acquisitions: a) standard apical view, b) combined apical view. Notice how myocardium walls are more visible on combined view.

From another patient, Figure 6.21 shows another two apical windows, a standard v_0 and, in this case, an interior displaced apical view, v_1. It can be seen how this second view, due to its inclination, has a better contrast on the mid posterior and posterior-lateral walls, as better appreciated in the short-view axis. Also notice how the delineation of the LV for the apical lateral segment is better on v_0, and for the mid lateral is better on v_1. Figure 6.22 shows the combination of these acquisitions.

Figures 6.23 to 6.25 show the temporal dynamics of heart walls. Note how the lateral wall contrast varies during the cardiac cycle, due to the different angles of insonification.

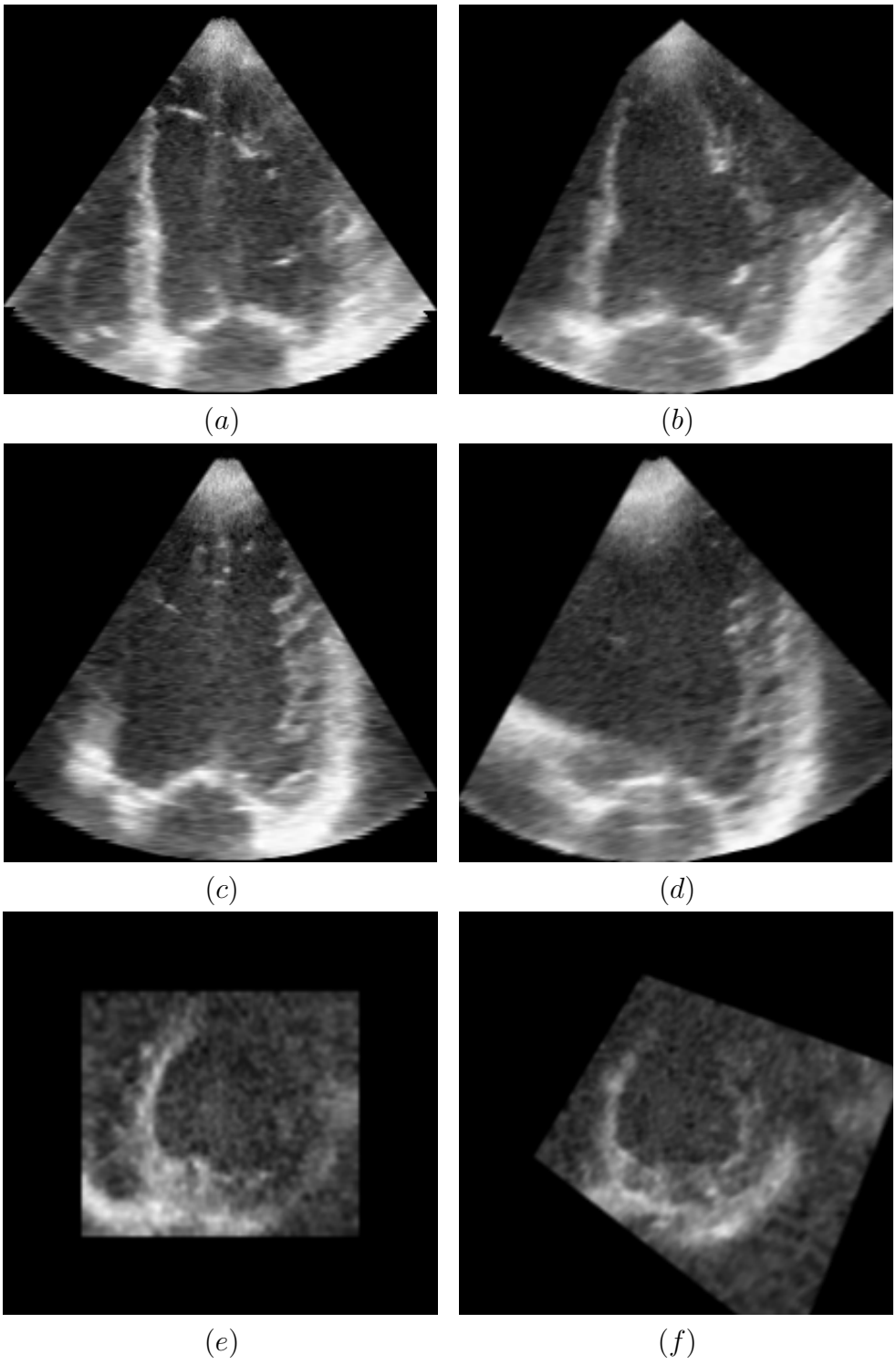


Figure 6.21: Original apical views. Left, v_0 from top to bottom: a) four-chamber view, c) two chamber view, e) short-axis view. Right, v_1 from top to bottom: b) four-chamber view, d) two chamber view, f) short-axis view.

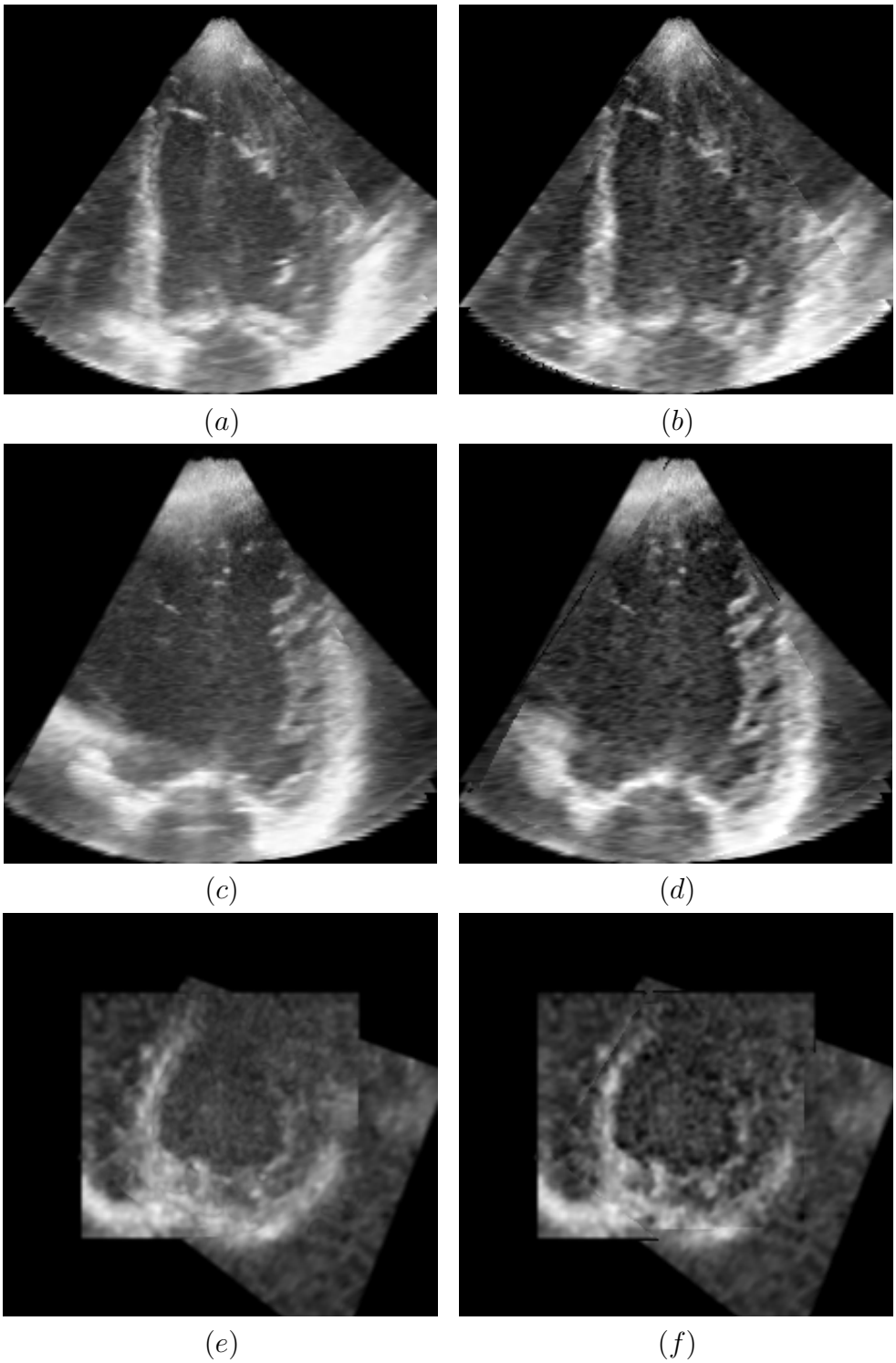
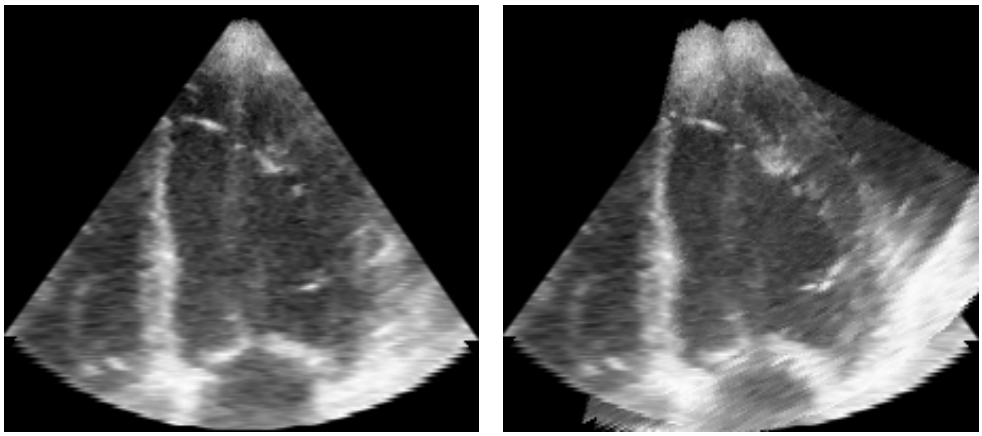
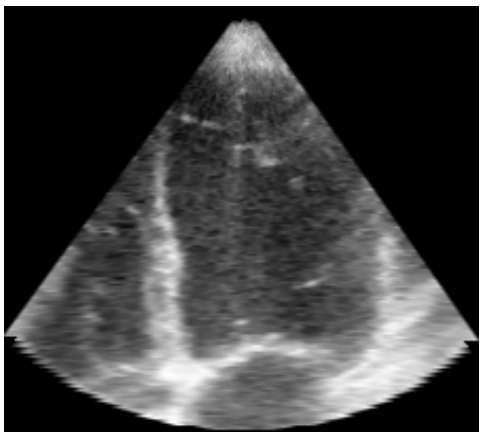


Figure 6.22: Combined apical views. Left, Maximum from top to bottom: a) four-chamber view, c) two chamber view, e) short-axis view. Right, multiview deconvolution from top to bottom: b)four-chamber view, d) two chamber view, f) short-axis view.

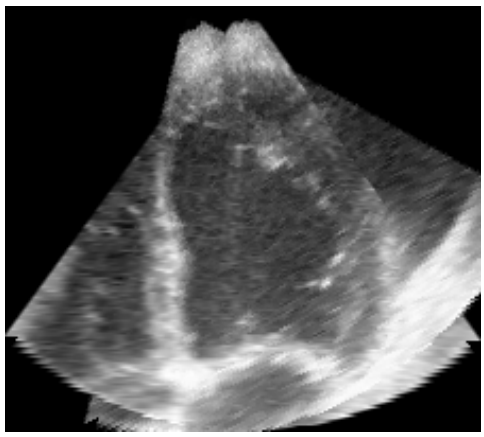


(a) $0ms$

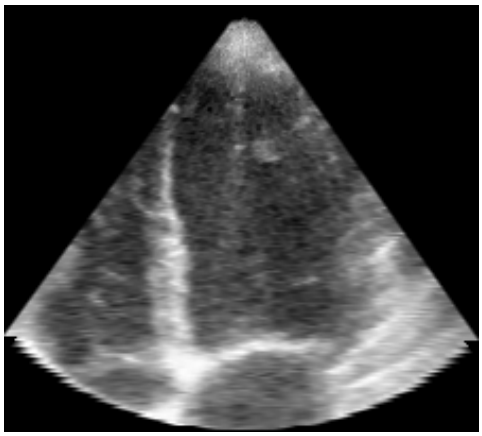
(b) $0ms$



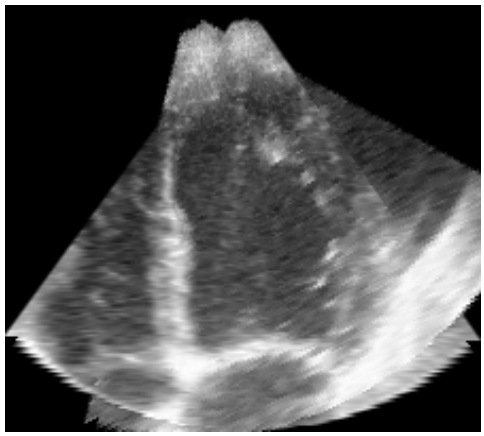
(c) $100ms$



(d) $100ms$

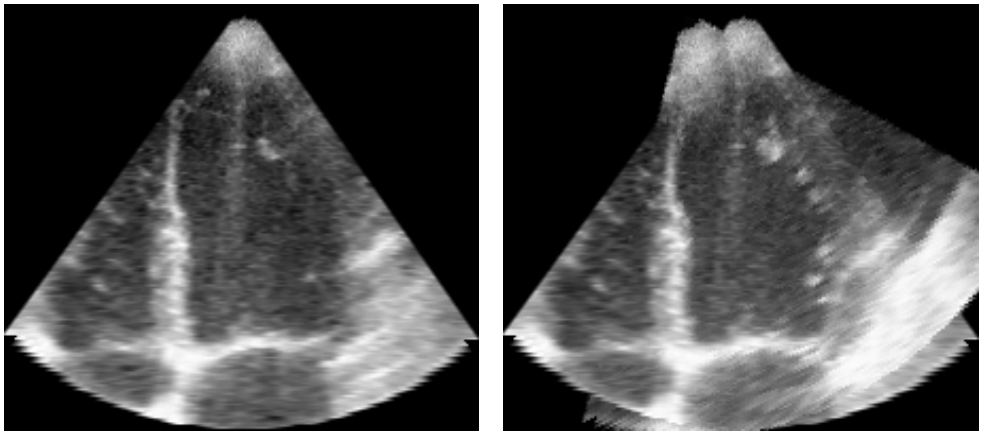


(e) $200ms$

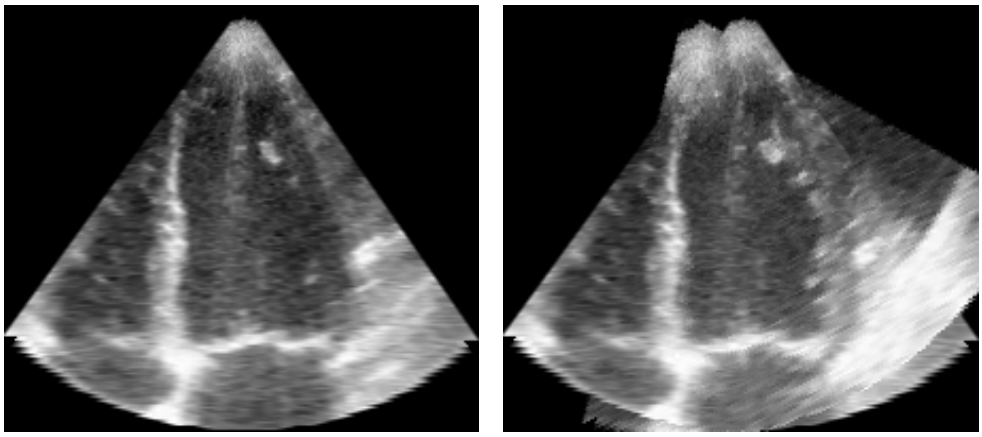


(f) $200ms$

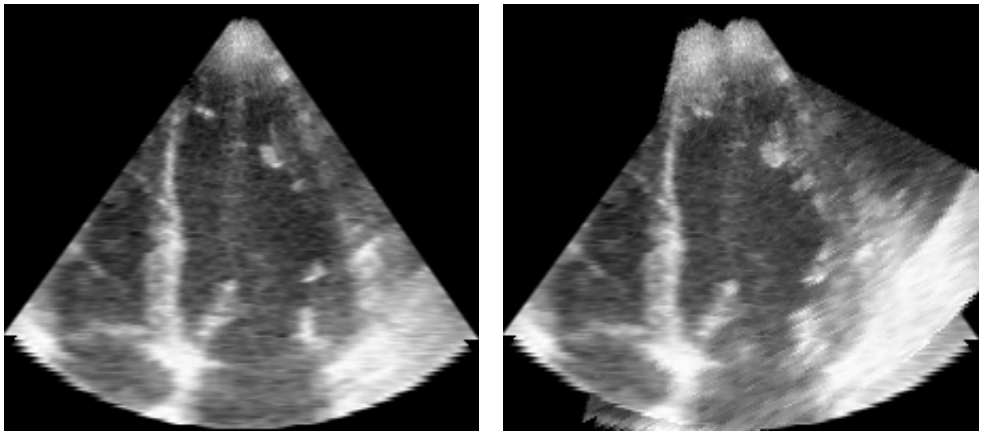
Figure 6.23: Four-chamber view of apical acquisitions at different frames of the cardiac cycle. Left, standard apical. Right, combination with external apical.



(g) 300ms (h) 300ms

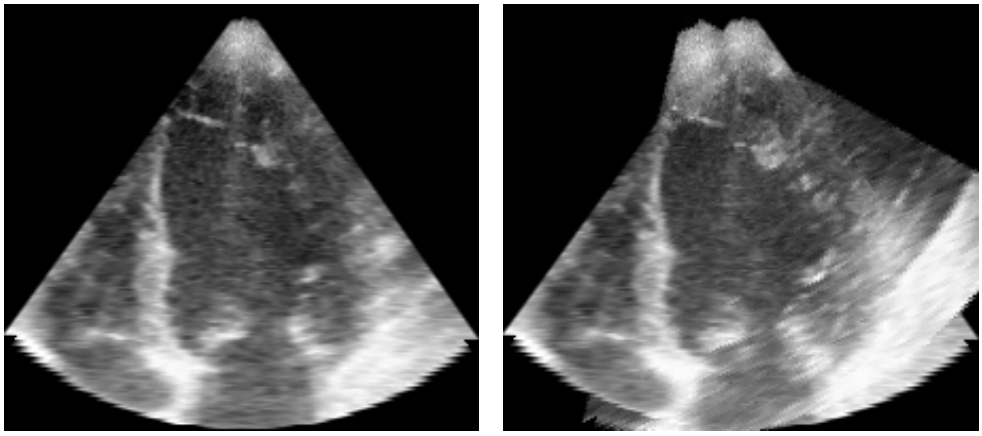


(i) 400ms (j) 400ms

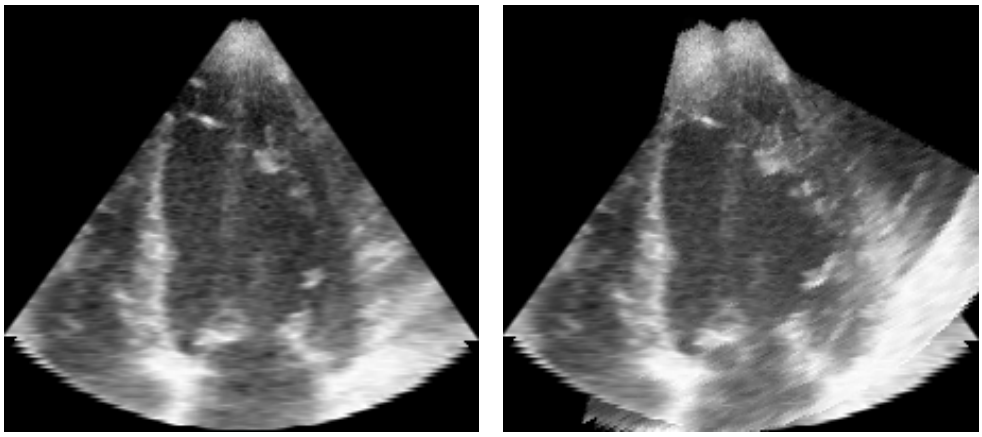


(k) 500ms (l) 500ms

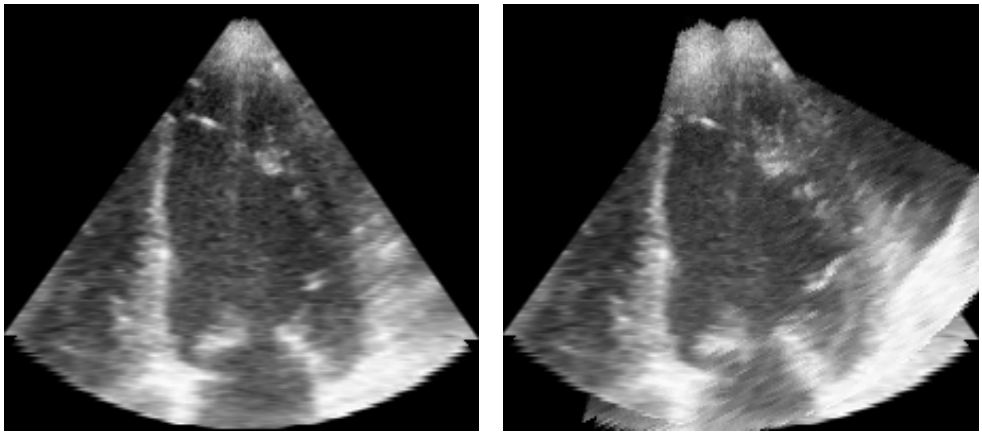
Figure 6.24: (*cont.*) Four-chamber view of apical acquisitions at different frames of the cardiac cycle. Left, standard apical. Right, combination with external apical.



(m) 600ms (n) 600ms



(o) 700ms (p) 700ms



(q) 800ms (r) 800ms

Figure 6.25: (*cont.*) Four-chamber view of apical acquisitions at different frames of the cardiac cycle. Left, standard apical. Right, combination with external apical.

6.5.3 Apical-Parasternal Acquisitions

Potentially, views with more different positioning are more interesting to combine. In some patients the apical window shows a clearer view of the heart walls, while in others, it is the parasternal window which shows better results. These two windows are very complementary in the sense that the ultrasound beam follows almost a perpendicular trajectory, therefore the walls that are most visible in one view are least visible in the other one and viceversa. Indeed, patients may have different image quality in apical and parasternal windows, and typically, both views are acquired.

Figure 6.26 shows the original acquisitions from apical and parasternal windows. It can be noticed that in difference to the apical acquisition, the parasternal one shows a better image quality in general, particularly noticeable on tissue delineation and wall contrast, e.g. in the lateral wall. However, the parasternal acquisition lacks some important areas, such as the apex, which will be hardly ever captured given the position of the probe and the angle of vision of current ultrasound probes. This information is present in the apical acquisition.

Initial results with apical to parasternal experiments are presented in this section. As visible on the images, some registration issues are still present. Those were registered only with the mesh-based approach, since the automatic procedure did not converge. One possible explanation is the small overlap between the two data sets, even after the mesh-based initialization as explained in Section 6.4.3. Note, however, that the differences between the initialization meshes in the apical and parasternal views may be also larger than in the case of only apical displaced views. This fact makes that the automatic algorithm starts further away from the optimal position and is not able to reach it.

Figure 6.27 shows the fusion of these two acquisitions. We can notice the errors in misalignment in the area of the mitral valve. However, the apical segments show a good continuity, especially the septal and lateral walls. Given the difference of the two data sets, the maximum operator tends to create a rather confusing image with many bright elements. The multichannel deconvolution makes walls thinner, especially on the apical view, as visible in the mid and basal lateral segments. This result, along with the ones seen in the multiple apical section, suggests that the technique works best when the acquisitions have an accentuated different point spread function. In comparison to simpler techniques, the diversity between the different views makes a larger difference on the results.

Figure 6.28 shows a simple experiment where the parasternal view is integrally kept, and is complemented outside its domain the apical view is, such as the apical region. In this way, the quantification of the left ventricle volume may be feasible using the better border definition of the parasternal window. Although this is a simple technique, which needs only the registration step, this complementary of data can be benefitial for clinical diagnosis purposes.

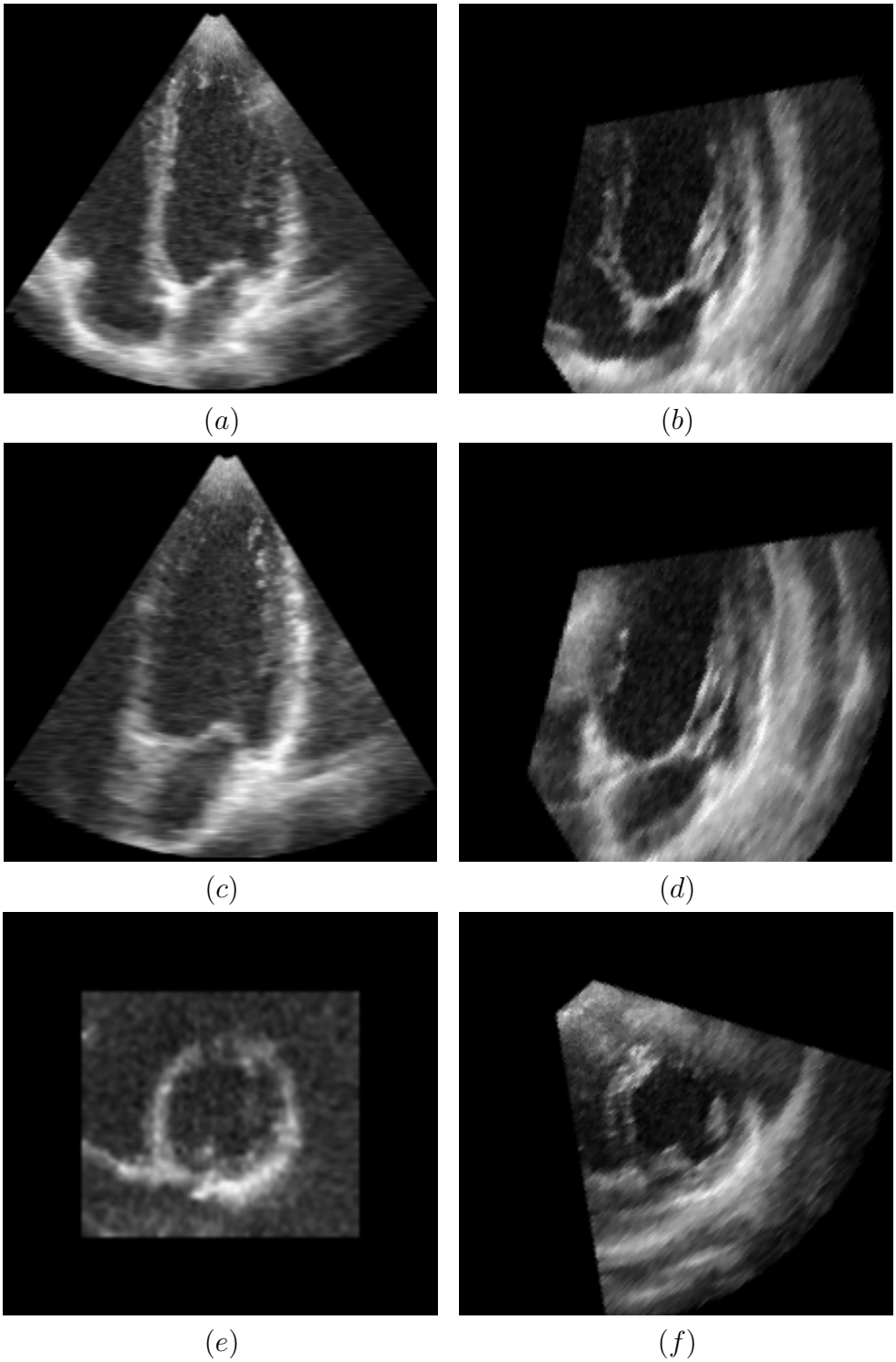


Figure 6.26: Original apical and registered parasternal view. Left, apical, from top to bottom: a) four-chamber view, c) two chamber view, e) short-axis view. Right, registered parasternal, from top to bottom: b) four-chamber view, d) two chamber view, f) short-axis view.

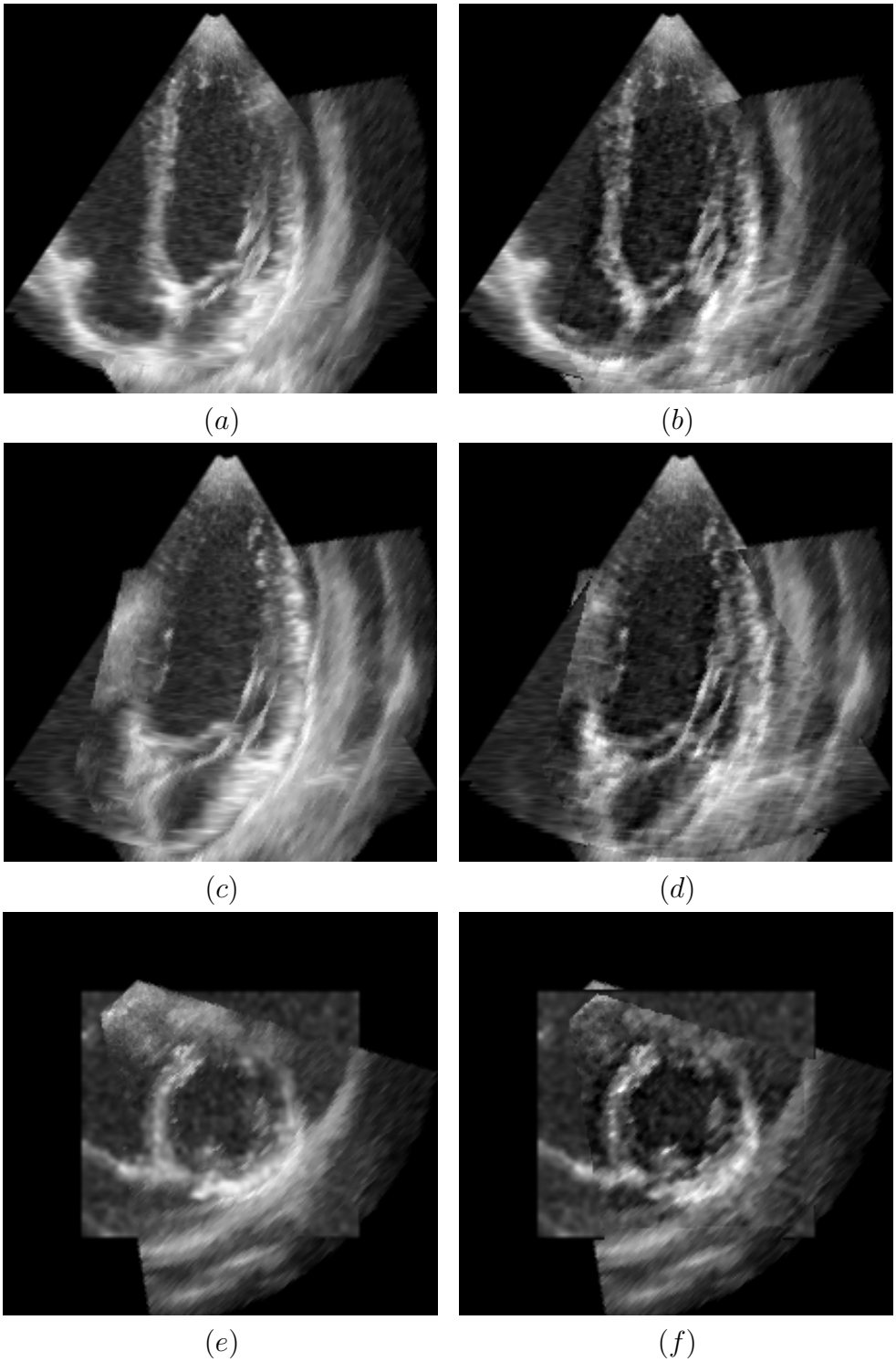
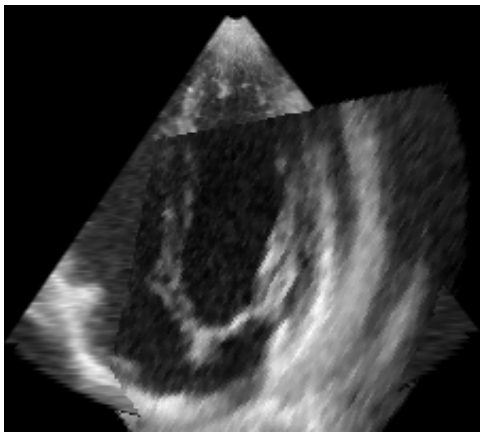
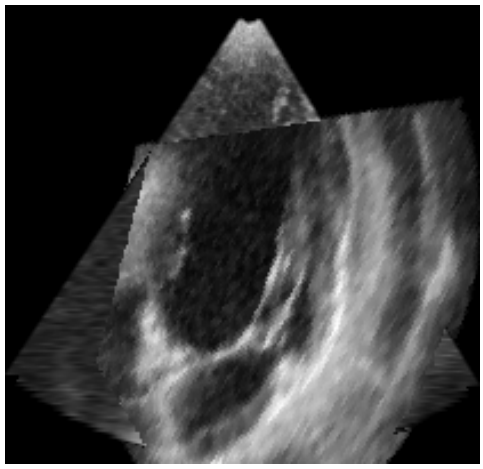


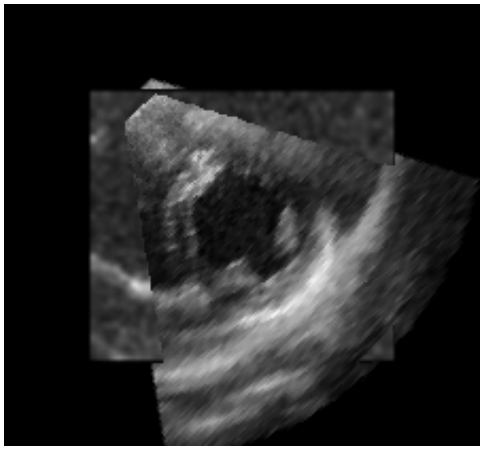
Figure 6.27: Maximum and multiview deconvolution of apical and registered parasternal view. Left, maximum technique, from top to bottom: a) four-chamber view, c) two chamber view, e) short-axis view. Right, multiview deconvolution, from top to bottom: b) four-chamber view, d) two chamber view, f) short-axis view.



(a)



(b)



(c)

Figure 6.28: Parasternal view complemented outside its domain with the apical view: a) four-chamber view, b) two chamber view, c) short-axis view.

6.6 Conclusions

We developed and implemented a technique to combine acquisitions from the heart from different acoustic windows, with the aim to increase the field of view, the visibility of the walls, reduce noise and improve the spatial resolution. The first two goals were notably achieved. From the clinical perspective, this is interesting because it enables examining the whole heart in a single dataset, which is of special importance in dilated hearts. Indeed, the presented fusion techniques may permit improving the performance of automatic quantification algorithms.

Spatial resolution, which showed to be best in multichannel deconvolution, did not show the same degree of improvement as in Chapter 5 for the breast application. The requirements for cardiac applications, the limited spatial resolution of the original acquisition, the free-hand nature of the acquisition that limits the registration between the different views and the lack of diversity between the different views are the causes of this fact. Initial tests for the combination between parasternal and apical windows show how multiview restoration can help combining those data sets, although the poor quality of registration limits the performance of the restoration algorithms. Noise within cavity was reduced in some techniques, although this showed to have less relevance for clinical practice than border delineation.

There are additional limitations of the system. It relies on the periodicity of the cardiac cycle, which showed to be a reasonable hypothesis in most of experiments. For a more general application of the technology, a temporal alignment within the cardiac cycle should be envisaged. Indeed, as conceived, the system does not work in real time, in the sense that it works with stored data. If the registration process could be sped up, it may be possible to store a first view and complement it with other views in real time, as the current panoramic systems in two-dimensions. At last, the fusion of the different views unfortunately also fuses some artifacts, such as clutter. These factors should be considered in future investigations. Moreover, the availability of several views could indeed help removing these artifacts out.

Further work includes improving the registration robustness. For instance, the metric could include the saliency measure in order to weight the most important areas to be registered. The need of a reference mesh for initialization is somehow a limiting factor we encountered, since it is not always possible to obtain a reliable segmentation in arbitrary views. Also the temporal dimension could be integrated into the registration step in order to make it more robust. Indeed, the temporal diversity could also help the fusion process to denoise, as in temporal compounding techniques (e.g. [15]). Other improvements include model-based fusion techniques, such as using an implicit model of walls to ensure their continuity, capturing elements of the wall both in the spatial and temporal dimensions. Elements of further work include ongoing validation with dilated heart patients.

CHAPTER 7

Conclusions and Future Work

In this thesis, we tackled the problem of restoring ultrasound echographic imaging from different acquisitions. This approach has shown great success in 2D imaging in the recent years with simply averaging the different views. We developed several new advanced techniques that, in addition to improve the signal-to-noise ratio, are able to improve the spatial resolution. This is particularly important in the addressed 3D applications, due to the anisotropy of the point spread function of the system.

We introduced the concept of *multiview* deconvolution. To develop it, we first reviewed the deconvolution methods to understand the limitations of inverting the degradation of a linear system. In general lines, frequencies eliminated by the imaging system are not recovered, needing a regularized solution to cope with the ill-posedness of the problem. We then introduced the generalization to multichannel deconvolution, where several acquisitions are available. In this case, the ill-posedness of the problem is reduced, since each view restricts the space of solutions. In the particular case where the different acquisitions are done with the same system but from different positions, which we denote as multiview, each acquisition permits to recover a part of the spectrum. Each view contains the features in a certain orientation, and the presented techniques permit to combine them into a single volume.

To estimate the PSF, we focused on multichannel PSF estimation techniques, particularly on subspace techniques. While theoretically the blind multichannel deconvolution problem is overdetermined and the PSF and the original volume can be obtained from the data channels only, in practice the system is very sensitive to noise and the *a priori* knowledge of the size of PSF. At this point, we introduced the multiview constraint, where the solution space is reduced by adding some *a priori* constraints about the geometrical relationship between the different views. However, this constraint does not guarantee the performance of the system within reasonable limits for real applications. Therefore we introduce realistic constraints about the shape of the PSF (three-dimensional Gaussian) to further increase the robustness to noise and support. We verified that the hypothesis of Gaussianity is well adapted for our acquisitions, obtaining a satisfying estimation of the PSF.

Deconvolution methods had been priorly developed for ultrasound systems, but specially in the domain of the RF signals. In the domain of the envelope signals, the most important problem was to not enhance speckle edges. We solve this issue by the

use in the deconvolution process of several acquisitions which have uncorrelated speckle patterning, and with the use of edge-preserving regularization terms. In this way, strong reflectors coherent with the system degradation are enhanced, while random patterning is attenuated.

At a second stage, we introduced another approach to combine the different views. This second approach, which we denominate *multiview fusion*, consists in building a fused volume by detecting the features of interest on each of the acquisitions. To this goal, we propose two new techniques in the spatial domain, weighted and generalized averaging, and two new techniques in a transformed domain, maximum frequency and wavelet fusion.

- *Weighted Averaging* consists in assigning a different weight to each view, locally, in order to capture the features of interest. We defined a saliency measure, according to each application, and filtered by a speckle detector in order not to consider speckle patterning. The distinction of speckle from features is not always possible. However, the fact of using multiple views makes this well-known problem easier, in the sense that it is not the absolute presence of features that matters but its relative presence. In practice this works relatively well, better preserving features of interest than simple averaging. However, weighted averaging does not guarantee the preservation of the best spatial resolution.

- We introduced *generalized averaging* in order to keep the advantages of weighted averaging while better preserving the spatial resolution. It is based on the discrepancy of the data sets, switching from the minimum operator to averaging to the maximum by the generalized averaging. The idea is to use the minimum when data are consistent, average in the case of speckle and maximum in case of inconsistency. Although it allows to obtain better results on synthetic data, the data discrepancy metric is difficult to be obtained. Maybe it would be more desirable to compute the min, average and max operators and combine them in a more controllable fashion (e.g. weighted average) than generalized average.

- In the transformed domain, we proposed the maximum frequency technique. Given the fact that each view contains a sector of the spectrum, the idea behind this technique is to build a fused volume by those coefficients of maximum amplitude. This technique does not depend on any parameter and therefore does not need any kind of estimation. Indeed, it succeeds to gather most of different features from each acquisition. However, all noise is also maintained, improving to a smaller extent the signal-to-noise ratio than other techniques.

- This research line evolved to use a wavelet approach with the same objectives but with a better spatio-frequential control. This permitted to avoid some low-frequency artefacts present with the maximum frequency technique. Indeed, the wavelet decomposition enabled to introduce coefficient shrinking algorithms in order to control the amount of noise present in the final fused volume.

Tests on synthetic data showed the clear superiority of blind multiview deconvolution techniques over the fusion techniques. The reason is that deconvolution methods

revert the degradation of the system, while fusion techniques only preserve the different information content of each view. Within the fusion techniques, the different spatial domain techniques showed a similar behavior, significantly better than averaging. While feature preservation is achieved by all techniques, generalized averaging and spectral techniques show a slightly better preservation of spatial resolution. However, in some cases this is at the price of reducing slightly less the signal-to-noise figure. To solve this, the denoising stage available with wavelet fusion offers an interesting approach.

We used all this novel techniques in two new applications, 3D ultrasound breast and cardiac imaging. These applications are a technological breakthrough on their own, and the multiview restoration techniques represent even an improvement of their performance. Alternatively, the mutliview techniques could be an opportunity to reduce the number of acquisitions, thus the time of examination.

For breast imaging, the most important goal was to overcome the anisotropic resolution of linear arrays. By obtaining scans at different angles, this goal was achieved with all techniques, including averaging. Blind multiview deconvolution techniques showed the best quantitative results on phantom data and probably the best results in *in vivo* data, improving: tissue delineation, which permits pathologic boundary discrimination; contrast resolution, which helps in discriminating masses; and spatial resolution, which may potentially help in diagnosing microcalfications. We also studied the speckle model on 3D log-compressed ultrasound images, which was found to be close to additive colored Gaussian noise, and the shape of the PSF, which was well described by a Gaussian shape.

For cardiac imaging, the most important benefits are to improve the field of view and the tissue contrast. Enlarging the field of view is important for patients with dilated hearts, which are typical cases for some cardiomyopathies. Heart walls have typically some missing segments, which difficult their analysis both qualitatively and quantitatively. Improving the tissue contrast improves their analysis. A registration method of the different views was developed. With the help of a manual initialization, it can successfully put different apical acquisitions in correspondence. Indeed, multiview restoration techniques improved endocardium delineation and the signal-to-noise ratio. Deconvolution techniques showed an increase of spatial resolution by the combination of different apical windows, although it was not very significant. Initial studies with the combination of apical and parasternal indicate that with acquisitions with PSF that differ more, spatial resolution can be further improved.

Many extensions and improvements arise from this work. For the PSF estimation, the subspace techniques showed to be sensitive to noise and support estimation. Future work includes comparing our approach to other estimation techniques, notably alternate maximization (AM) [141], where the point spread function is estimated and the image is restored alternatively. Other techniques, maybe not based on the channel disparity, could be envisaged. Particularly, given the colored nature of speckle, estimate the PSF from the autocovariance of homogeneous areas would be in principle

possible [79].

Other authors propose to log-decompress the images in order to use more accurate speckle models, although this presents some problems. For instance, the need to know the constants involved in the process and the quantization effects due to the non availability of the 16 bit data from the machine. Regarding the regularization techniques, these models could be introduced in this term. Moreover, it could be envisaged to consider the multiple views within the regularization term rather than the reconstructed volume only.

The treatment of the raw scan lines rather than the scan conversed ones could improve the performance of the algorithm, since no interpolation effects would be present in the treated images. Indeed, for the cardiac application, due to the sparsity of the data, a great improvement could be obtained by during the scan conversion from the two (or more) datasets into the cartesian axes. The temporal dimension could also be taken into account for denoising purposes [15].

The multiview techniques could be extended to other applications, including obstetrics, where the market is more and more demanding on high quality images and liver imaging, where there is a need for an extended field of view. They could also be extended to other ultrasound modalities, such as frequency compounding, in order to better exploit the point spread function diversity present in the different acquisitions, and even to other fields, such as synthetic aperture radar, where different looks are also obtained to reduce the speckle patterning.

APPENDIX A

Ultrasound Echography

In this appendix we present the basic principles of ultrasound echography imaging for medical purposes. Many excellent references can be found on this subject, such as [167].

A.1 Medical Ultrasound Imaging

Ultrasound echograhy imaging is one of the most widely used since it permits to observe all kind of soft tissues in a non-invasive way. It is particularly useful for delineating the interfaces between solid and fluid-filled spaces. It permits to obtain real-time images, where the operator can select the most useful section for diagnosis and set the imaging parameters during the exam, often enabling rapid diagnoses. It shows the structure as well as some aspects of the function of organs. It is used for almost all soft tissue organs, including: heart, kidney, liver, vascular system and neonatal brain. In more than 50 years of use, no long-term side effects have been reported, and rarely causes any discomfort to the patient. Equipment is available almost everywhere; easy to carry, examinations can be performed at the patient's room. Indeed, it is relatively inexpensive compared to other medical imaging modalities, such as X-ray or magnetic resonance imaging (MRI). Given its real-time and non-invasive characteristics, it is widely used during surgical operations, often in combination with other imaging modalities such as pre-operative MRI.

However, ultrasound shows several limitations. Ultrasound waves can not propagate trough excessive impedance changes, such as penetrating bones, nor gases, like air in the lungs. This limits their application to "protected" organs such as the brain, or, to a less extent, the heart. Even in the absence of bone or air, the depth penetration of ultrasound is limited, making it difficult to image structures that are far away from the probe. This point is critical for obese patients, for instance. In fact, the image quality is very much dependant on the patient. And also on the practitioner. A high level of skill and experience is needed to acquire good-quality images and make accurate diagnoses.

We list the most common ultrasound imaging modalities used in medical practice. Figure A.1 shows an example of an ultrasound exam.

Figure A.1: Ultrasound echography is real-time, non-invasive, portable and relatively inexpensive medical imaging modality [6].

A.1.1 B-mode

B-mode imaging refers to 2D ultrasound images. This imaging modality will be described in Section A.2. Figure A.2 shows an exam of the breast tissue.

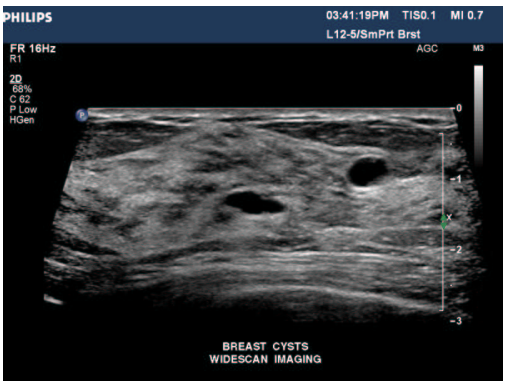


Figure A.2: Breast-tissue B-mode ultrasound echography exam [6].

A.1.2 3D

Recently introduced into the market, real-time three-dimensional (RT3D) echocardiography systems use a fully sampled matrix array to image volumes in a single acquisition. This allows for precise quantification of masses volume, since no hypothesis must be made from single slices of data. Currently, the applications which take more profit of this invention are obstetrics and cardiac applications. Figure A.3 shows an exemple of the use of the 3D technology for an obstetric application.

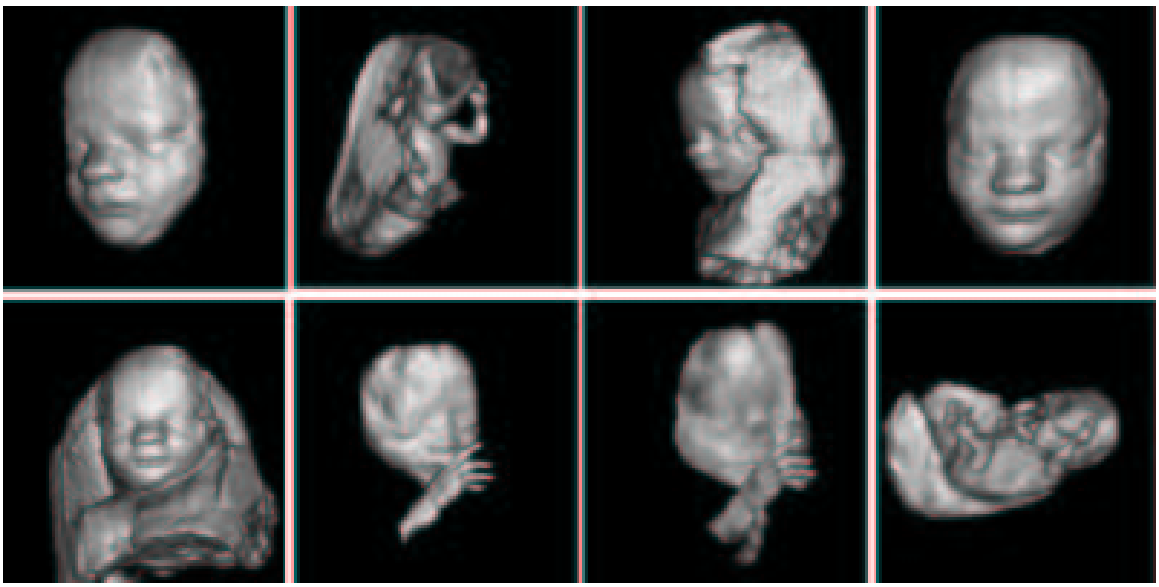


Figure A.3: Three-dimensional obstetric imaging [6].

A.1.3 Harmonic Imaging

Harmonic imaging consists in exploiting the non-linearities of the acoustic wave propagation. As the fundamental frequency wave is propagates, it is distorted creating harmonics. These harmonics are created at the tissue itself, showing the following benefits with respect to the fundamental frequency:

- higher frequency (typically x2, the second harmonic), thus better spatial resolution;
- lower attenuation, since it only travels from the tissue to the probe;
- less reverberation;
- smaller side lobes.

Areas of the image that are far away from the probe benefit most of harmonic imaging, due to the fact that the wave needs to propagate for a certain distance to get distorted.

Therefore, at the generation point they are less attenuated and less affected by phase-aberration. However, at reception the second harmonic has to be filtered from the fundamental wave signal, which may reduce the effectiveness of harmonic imaging.

There exist other techniques which take profit of the non linearity of the medium, such as pulse inversion [85]. This technique consists in sending two consecutive ultrasound pulses, the second pulse being an inverted copy of the first one. Linear scattering of the two pulses will give two echoes which are inverted copies of each other, and these echoes will therefore cancel out when added at reception. On the other hand, non-linear scattering, e.g. from bubbles, will give different responses. The harmonic components add, and the signal intensity difference between non-linear and linear scatterers is therefore increased.

A.1.4 Doppler

One of the benefits of ultrasound imaging is the ability to measure the velocity and direction of flows and tissues. The Doppler principle states that the reflected frequency changes proportionally to velocity of the imaged body [48]. Therefore, velocities of the blood flow will be detected as changes in the ultrasound wave frequency. High-amplitude, low frequency signals from moving tissues are typically filtered away to detect blood flows.

Two different kinds of Doppler techniques are available:

- Continuous Wave Doppler corresponds to the direct application of the Doppler principle with a sinusoidal wave. It is rarely used since it does not provide information about the position of the target.
- Pulsed Doppler evaluates the movement of the tissue by analyzing the differences of different acoustic pulses. The user must define a region where the Doppler image has to be computed.

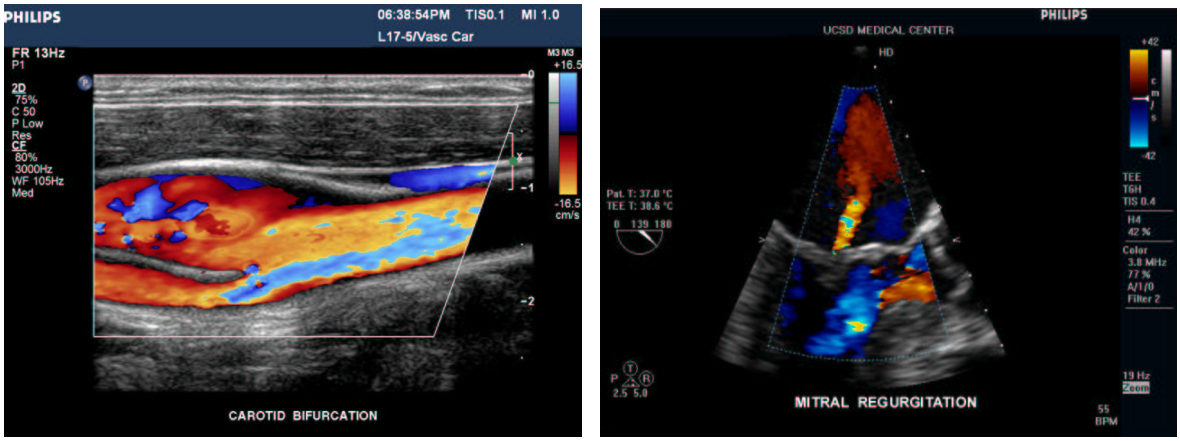
When Doppler information is displayed as colored images over the regular B-mode echography, it is named as *color* Doppler. Figure A.4 shows two examples of this modality.

Power Doppler [106], or Spectral Doppler, refers to the display of the power of the Doppler signal instead of the frequency shift. In this way, the system is sensitive to the amount of flow rather than its direction. In comparison to color doppler, power doppler has higher sensitivity to flow, better edge definition and depiction of continuity of flow.

Tissue Doppler imaging [42] consists in filtering the low amplitude, high frequency signal to detect the tissue movements such as myocardial contractions.

A.1.5 Stress Echocardiography

This modality of echocardiography consists in evaluating the heart in stress conditions. In this way, it might be easier to detect parts of the heart that may not be receiving enough blood or oxygen because of blocked arteries. Stress can be induced either via medicals (e.g. dobutamine) or via physical exercise (e.g. static bicycle).



(a) (b)

Figure A.4: Examples of color Doppler imaging: a) blood flow turbulences in the carotid vein, b) mitral regurgitation [6].

A.1.6 Contrast Agents

Cardiac ultrasound images can be improved by the use of contrast agents. Those are typically tiny air bubbles (less than 10 μm) in a biodegradable shell, that are introduced in the blood torrent. Due to the different acoustical impedance of air and blood, these elements greatly increase the contrast of the ultrasound image, permitting a better discrimination of the blood pool and the myocardium wall, the study of perfusion and capillary blood-flow.

Due to the non-linear reflectivity of the bubbles, harmonic imaging takes importance with contrast agents. The second and further harmonics are greatly reflected by the microbubbles, while the tissue does not.

A.1.7 Elastography

Cancer tissue elasticity can be used for discrimination from other tissues such as fat or benign lesions. Elastography [118] consists in creating a map of the elasticity of the tissue. There are manily two approaches:

- quasi-static [117], where a compression is appied and the resulting components of the strain tensor estimated via autocorrelation,
- low-frequency vibrational (50 - 500 Hz) [54], where the movement of tissues is evaluated with Doppler imaging. The vibration can be induced manually by the practioner, by the patient, with external vibrators attached to the body or with acoustic transducers.

Due to the elasticity of non-pathologic tissues, these move under the vibration and appear colored. On the contrary, malignant tissue remains uncolored since they are firm.

Recently, Bercoff *et al.* [22] propose the supersonic shear imaging technique which consists in creating focused shear waves with the ultrasound probe and image those at high frame-rate. This technique permits to obtain the elasticity map of the tissue, thus determining the nature of the observed lesions.

A.2 Image Formation

A.2.1 Beamforming

In order to focus the energy at a certain focal distance, different delays are applied to each of the elements of the transducer. In this way, ultrasound waves sum their energy constructively at the focal point. This process is known as beamforming. Transducers that operate in this way are named phased arrays, for their ability to control the phase of the acoustic pulses of each of their elements.

The delay τ applied to each element is easily computed, from the distance from the transducer to the focal point $d(x)$, as:

$$\tau(x) = \frac{1}{c}d(x) = \frac{1}{c}\sqrt{(x_f - x)^2 + (y_f)^2} \tag{A.1}$$

where c is the speed of sound, x is the position in the linear transducer, (x_f, y_f) is the position of the focal point. The speed of sound is used in this computation; therefore, a misestimation of its value will lead to beam defocalization.

Similarly, the reflected waves arrive with different delays to the transducer, enabling their localization in space. Following the same law expressed in Equation (A.2.1), a different delay is applied to each element to detect the reflections at a certain focal point. The resulting signal is known as the RF signal.

Figure A.5 shows a diagram of the beamforming operation. Typically, transmission beamforming is performed at one to four focal zones, and reception beamforming is performed for all points of the image.

A.2.2 Propagation

Ultrasound imaging is based on the propagation of the acoustic waves through the human body. It allows the propagation of pressure wave, since it has elasticity and compressibility of a fluid [72]. In the context of linear acoustics, pressure wave fields are assumed to represent solutions of the linear wave equation. In one dimension, it can be expressed as:

$$\frac{\partial^2 p}{\partial z^2} = \rho_0 \kappa \frac{\partial^2 p}{\partial t^2} \tag{A.2}$$

where p is pressure, ρ_0 is the medium density at equilibrium, κ is its compressibility, and z and t are the spatial and temporal dimensions respectively. This is a reasonable approximation for ultrasound imaging, although as mentioned in Section A.1.3 non-linear effects actually happen in real tissue propagation.

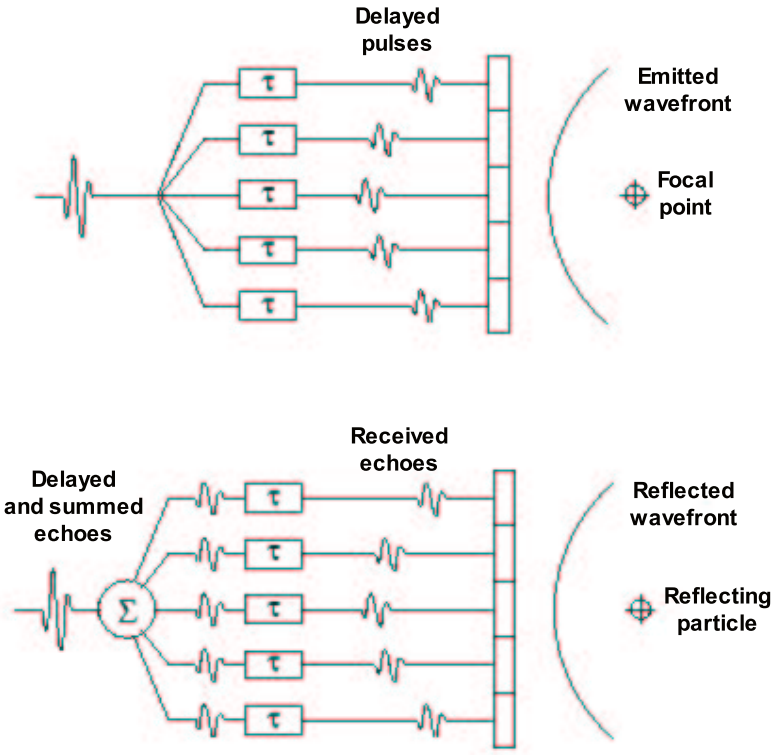


Figure A.5: Different delays are applied tp each transducer elements in order to vary the phase of the acoustic pulses. This operation is performed at emission (top) and at reception (bottom) [13].

A.2.3 Reflection and Scattering

Ultrasound echography is based on the echoes generated by the tissues from the incident acoustic wave. Reflection refers to the interaction of sound waves and particles which are much larger than the imaging wavelength; scattering refers to the interaction with particles which are comparable or smaller. The behavior is different. Reflection leads to specular echoes, which direction follows the Snell's laws of optics. Scattering consists in radiating in all directions. For particles much smaller than the wavelength, which are known as *Rayleigh* scatters, the radiation is isotropic.

A.2.4 Image Pipeline

Once the reflected echoes have been detected and beamformed, there is a series of operations done to accommodate this data for display. The envelope of the RF signal is taken, in order to avoid filter oscillations of the ultrasound wave and keep the envelope pulse. The time-gain-curve (TGC) is applied in order to compensate for the variation of attenuation with depth. Raw lines are converted to Cartesian data in the case of curved, sector or steered transducers. This signal is named as *envelope-detected* image. Due to the high range of echo intensities, images are log-compressed in order to be displayed in current monitors. This signal is named as the *log-compressed* image.

A.3 Performance Limits

A.3.1 Spatial Resolution

Spatial resolution refers to the ability of the system to discriminate closely spaced scatters. The shape of the point spread function (PSF), is determined by the Fraunhofer approximation. Briefly, this well-known expression from the optics literature states that the far-field complex amplitude pattern produced by a complex aperture amplitude function is approximately equal to 2-D Fourier transform of that function [58]:

$$PSF(\mathbf{x}) = \mathcal{F}\{A(\mathbf{x})\} \tag{A.3}$$

where $\mathcal{F}\{\bullet\}$ represents the Fourier transform, and $A(\mathbf{x})$ the aperture function.

Applied to ultrasound, this approximation states that the ultrasound beam pressure amplitude pattern can be estimated by the 2-D Fourier transform of the transducer aperture. In the lateral and elevational dimensions, this means that a rectangular aperture leads to a *sinc* function as PSF, with size inversely proportional to the size of the aperture. In other words, the bigger the aperture, the smaller the PSF. In order to avoid the secondary lobes of the *sinc* function, apodization functions are applied to the amplitude of the emitted pulses, such as the Hanning window [13].

The same rationale applies to the axial dimension, but in this case it is the shape of the ultrasound pulse which plays the role of the aperture. Typically a raised cosine modulates the ultrasound carrier. Both the frequency of the ultrasound wave and the length of the pulse will limit the spatial resolution on the axial dimension. However, these parameters are limited by the attenuation of tissues.

There is a trade-off between lateral resolution and depth-of-field (DOF). The DOF determines the length in the axial dimension at which the beam stays focused. It is [13]:

$$DOF = 8 f_{\#}^2 \lambda \tag{A.4}$$

where $f_{\#} = z/d$ (z is the depth, and d is the size of the aperture) and λ is the ultrasound wavelength. The larger the aperture, the smaller the DOF. In other words, for small PSF (large aperture), multiple focalization points will be needed to compensate the decrease of DOF. This will indeed impact on the frame rate.

Conventional ultrasound systems show a typical spatial resolution of about 300x1000x300 μm^3 in lateral, elevational and axial dimensions.

A.3.2 Speckle

The presence of scatterers smaller than the wavelength creates the so-called speckle noise. It is a random, deterministic, image patterning caused by the interference of the sub-resolution scatterer echoes. Its texture does not contain information about the underlaying tissue; its mean brightness does, in principle, reflect the brightness of the tissue.

Speckle has been widely studied in the literature. The pioneer works are from Goodman [58]. The first discussion of ultrasound speckle using statistical optics is in

Burckhardt [32]. A classic paper on the second-order statistics of speckle is by Wagner *et al.* [159]. A discussion of Rician statistics in tissue characterization is in Insana *et al.* [72]. We breifly describe these results in this section.

Assuming that the statistics of the RF signal are White Gaussian[58], that is, that the scatters are independent, the envelope-detected signal follow a Rice distribution

$$p(a) = \frac{a}{\sigma^2} \exp\left(-\frac{a^2 + s^2}{2\sigma^2}\right) I_0 \frac{as}{\sigma^2}, a \geq 0 \tag{A.5}$$

where a is the mean scatter spacing, s is the strength of a specular reflector, σ is the standard deviation of noise and I_0 is the incomplete Bessel function of order zero. The Rician PDF is parameterized by a single parameter k, which is defined as:

$$k = s/\sigma \tag{A.6}$$

The Rician distribution reduces to the Rayleigh distribution for the special case $s = 0$, that is, when there are no bright scatters. Then, the image consists in purely speckle patterning due to sub-resolution scatters. The Rayleigh distribution is defined as [159]:

$$p(a) = \frac{a}{\sigma^2} \exp\left(-\frac{a^2}{2\sigma^2}\right), a \geq 0 \tag{A.7}$$

However, log-compression changes the characteristics of the signal probability distribution. In particular, it affects the high intensity tail of the Rayleigh and Rician PDFs more than the low intensity part. Strictly, the log-compression on Rician PDFs is a Fisher-Tippett distribution [80, 45], expressed as:

$$p(a) = \frac{\exp(-z)z}{\beta} \tag{A.8a}$$

$$z = \exp\left[-\frac{a - \mu}{\beta}\right] \tag{A.8b}$$

However, some approximations can be made. If the PSF is considered as Gaussian, Kao *et al.* [79] demonstrate that speckle noise distribution can be approximated as a additive colored Gaussian PDF. Abd-Elmoniem *et al.* [9] assume that in a reasonably high number of scatters, that is, not too low SNR, additive white Gaussian noise is a good approximation. From our experiments, the additive colored Gaussian noise is the most satisfying model.

Many techniques have been proposed to reduce speckle, including spatial filtering [64, 88, 44, 61, 11], temporal integration [47], frequency compounding [40] and spatial compounding [155, 116].

A.3.3 Depth

High-frequencies suffer much more attenuation than low frequency ones. Table A.1 shows the attenuation values for different kinds of human tissue. The upper acoustic

power limit is bounded in order to not burn the tissues (FDA limit = 100mW/cm^2). Therefore, there is limitation in frequency which limits the spatial resolution and the presence of speckle patterning.

Typical frequencies are 3.5 to 10 MHz, for depths from 15 to 7.5 cm.

Type of Tissue	Attenuation ($dB/MHz \cdot cm$)
Liver	0.6 - 0.9
Kidney	0.8 - 1.0
Fat	0.5 - 1.0
Blood	0.17 - 0.24
Plasma	0.01
Bone	16.0 - 23.0

Table A.1: Attenuation values for different human tissues [77]

There exist high-resolution devices. They operate at frequencies form 20 to 50 MHz, and penetration depth go from about 10 mm to 5 mm respectively. Therefore, their application is limited to sub-cuteanous imaging and small animals.

A.3.4 Clutter

Clutter artifact is due to the reverberation of ultrasonic waves between tissue interfaces. Instead of a single clean echo, waves keep bouncing back and forth between the different interfaces. This is particularly visible near the edges of masses or other clearly definite interfaces such as inside the heart.

This artifact is specially important in Doppler imaging and in bubble specific imaging.

A.3.5 Shadows

Shadows occur below objects that reflect or absorb a greater portion of energy. Therefore, there is no transmitted energy and tissues that are further away are not insonified at all. This effect is completely blocking in the case of bones (e.g. ribs) or gas (e.g. air in the lungs). The difference of impedance between soft tissues and these elements is so large, that no energy at all is transmitted to tissues below them.

Shadows may be an indication of malignancy in breast cancer.

A.3.6 Drop-out

The reflected wave follows the Snell's law. Therefore, if the angle of the incidence of the emitted wave is perpendicular to the tissue, the reflected wave will also be. In this case, a strong echo will be received. On the contrary, when the tissue is almost parallel to the beam, most of the reflected energy will not return to the probe.

A.3.7 Field of View

The field of view depends on the size of the available acoustic window. According to this, different probes will be used. Transducers may be linear, sector or curved arrays. Linear array transducers produce rectangular images and offer the best overall image quality. Sector array transducers are appropriate to image larger organs from a small acoustic window, such as in between the ribs. Curved array transducers offer a compromise between these two kinds, and are optimal when the acoustic window is large.

APPENDIX B

Phantoms

In this appendix we briefly describe the composition of the phantoms used in our experiments.

B.1 ATS 539[5]

The Model 539 Multipurpose phantom is an easy, comprehensive means of evaluating imaging systems with an operating frequency range of 2.25 to 7.5 MHz. The phantom is designed with a combination of monofilament line targets for distance measurements and tissue mimicking target structures of varying sizes and contrasts. Cystic-like target structures are positioned in-line vertically, thereby permitting an entire target group to be displayed in one view. Due to the acoustic similarity of the background material and the target structures, artifacts caused by distortion, shadowing or enhancement have been eliminated. Six gray scale targets ranging in contrast from +15 to - 15 dB are provided to evaluate the system's displayed dynamic range and gray scale processing performance.

Figure B.1 shows a schematic illustration of the parts that contain this phantom.

B.2 Bubbles

In order to measure the point spread fucntion of the system a phantom was constructed with small air bubbles on a saline solution. No scatters are present within the saline solution. The diameter of the air bubbles is about 10 μ m.

This phantom was build by Rob Entrekin, from Philips Ultrasound Bothell, USA.

B.3 Grape Fruit

In order to measure the the spatial and contrast resolution of the three-dimensional breast imaging device, some grapes were suspended in gelatine.

This phantom was build by Rob Entrekin, from Philips Ultrasound Bothell, USA.

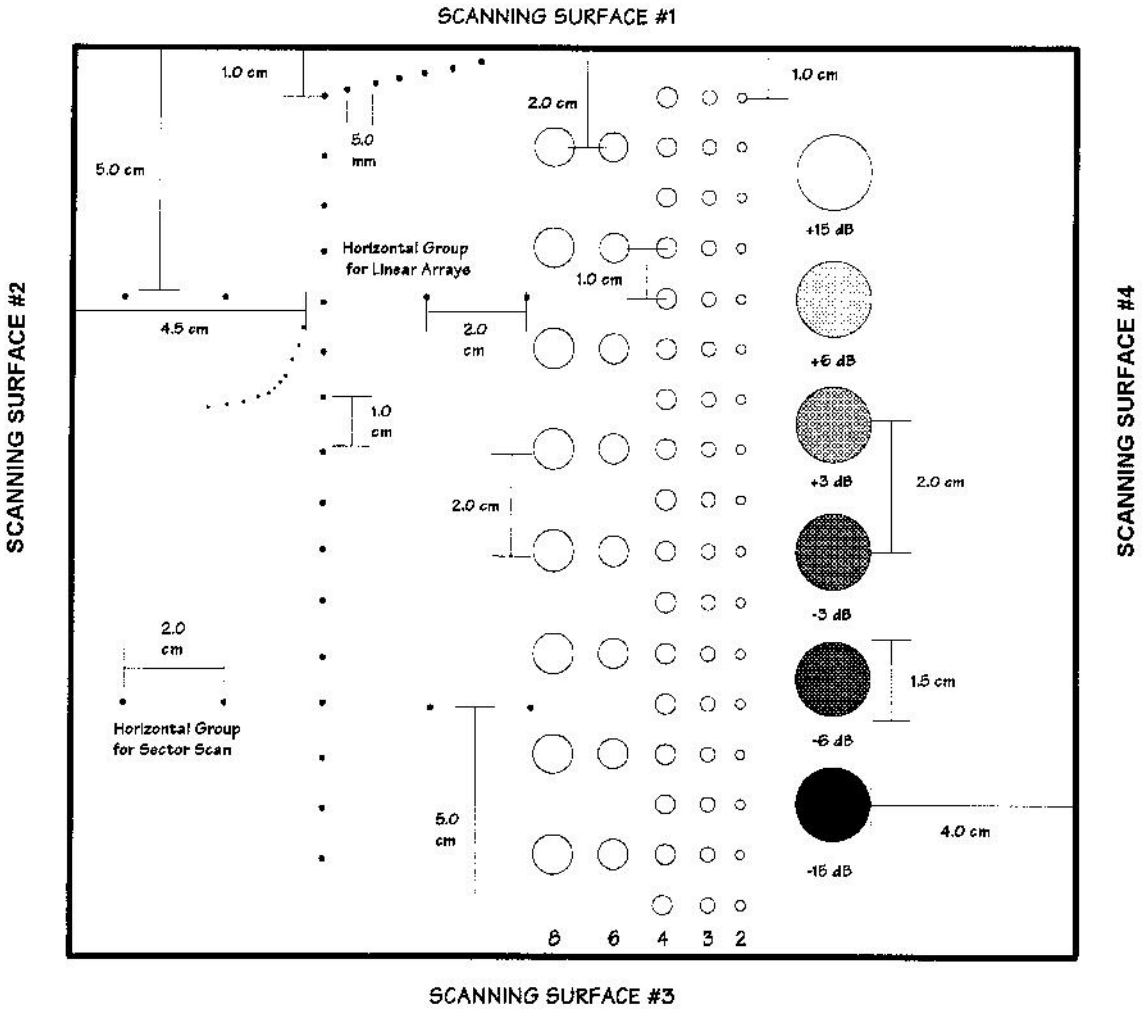


Figure B.1: ATS 539 phantom. Material: Urethane rubber, Sound speed: 1450 m/s, Absorption: 0.5 dB/cm/MHz. Left hand side: Line Targets (Monofilament Nylon, Diameter: 0.12 mm). Centre section: Anechoic Targets (Cylindrical in shape filled with homogenous (nonscattering) material. Diameter: 2 to 8 mm. Right hand side: Gray Scale Targets (Cylindrical in shape filled with scattering material. Diameter: 10 mm, Contrast is backscatter relative to background material.)

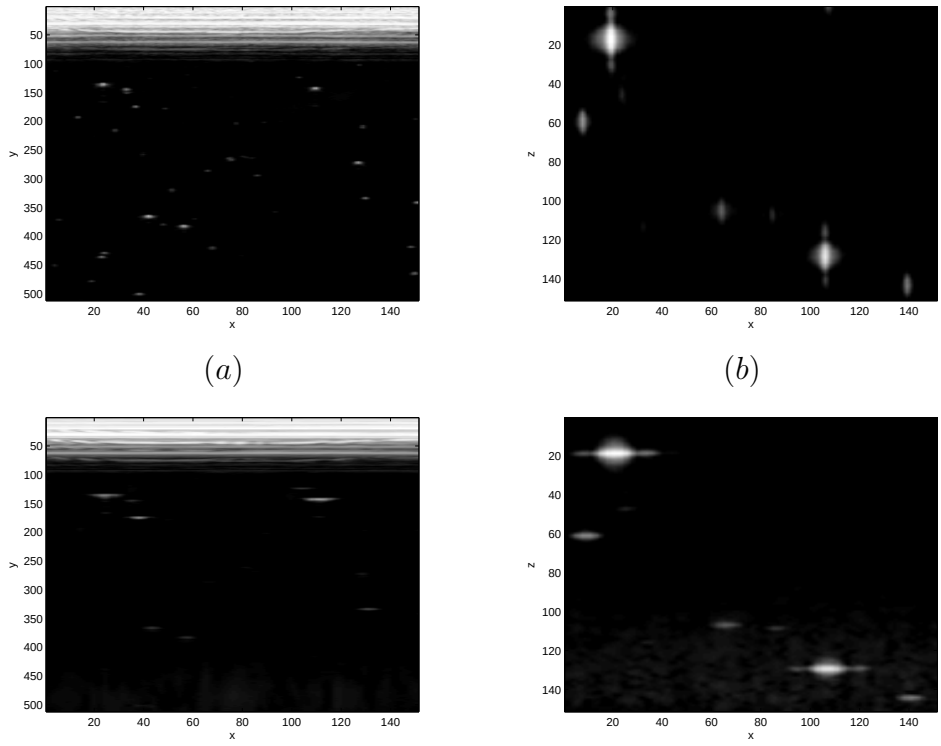


Figure B.2: Bubble phantom to measure the point spread function: a) B-scan($x = 80$) of v_0, b) C-scan($y = 400$) of v_0, c) E-scan($x = 80$) of v_{90}, d) C-scan($y = 400$) of v_{90}.

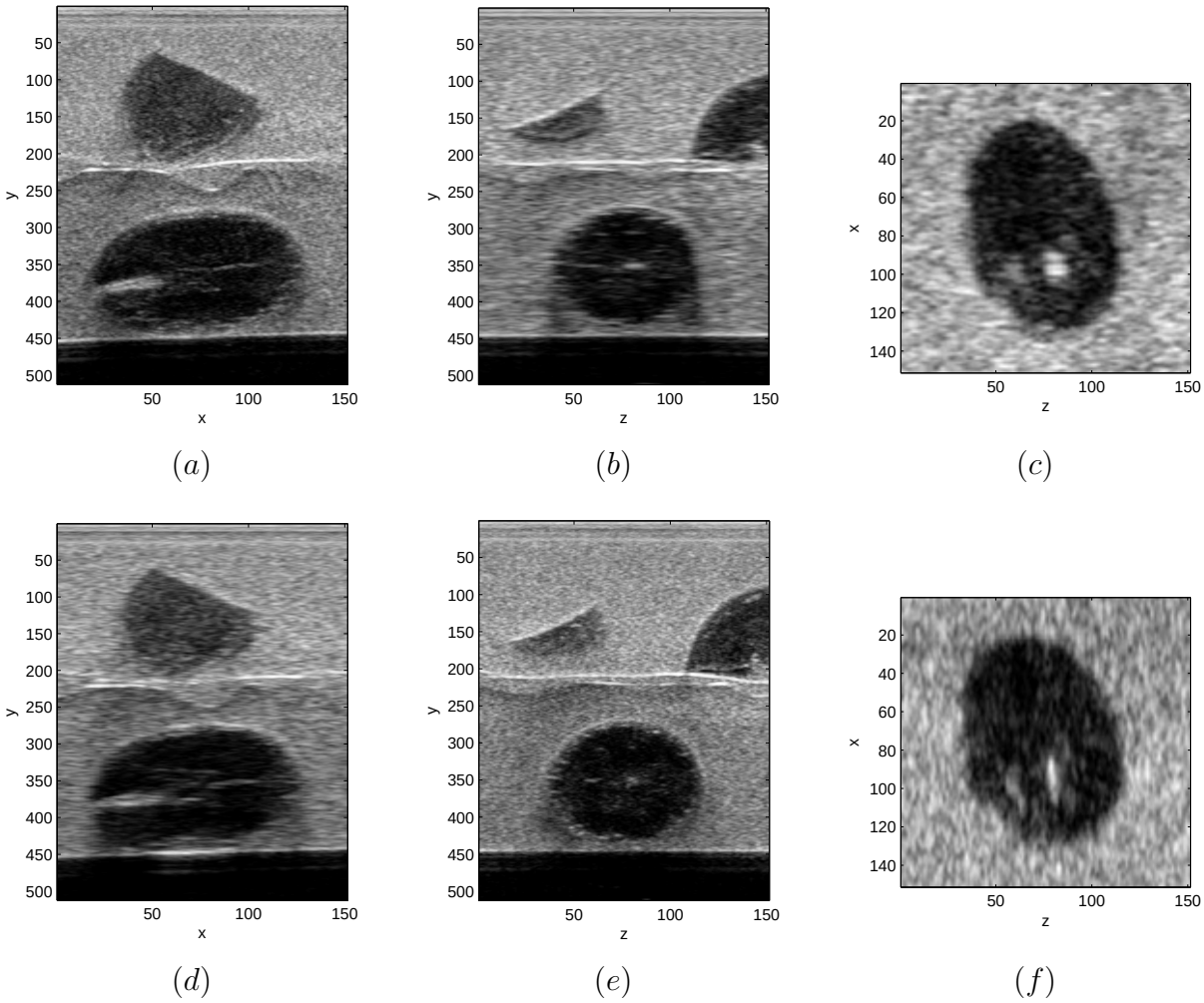


Figure B.3: Slices of fruit phantom 3D acquisition, zoom on a 3x5x3 cm area, voxel size 0.2x0.1x0.1 mm. For v_0 (upper row): a) B-scan ($z = 60$), b) E-scan ($x = 100$), c) C-scan ($y = 350$). For v_{90} (lower row): d) E-scan ($z = 60$), e) B-scan ($x = 100$), f) C-scan ($y = 350$). Due to the point spread function of the system, the best spatial resolution is achieved in the B-scan, which corresponds to the xy plane for v_0 and to the zy plane for v_{90}.

Bibliography

[1] (2005). American college of radiology. http://www.acr.org.

[2] (2005). European cardiovascular disease statistics. http://www.ehnheart.org.

[3] (2005). Heart centers online. http://healthcentersonline.com.

[4] (2005). National alliance of breast cancer organizations. http://www.nabco.org.

[5] (2005). Phantom ATS 539. http://www.atslaboratories.com.

[6] (2005). Philips medical systems. http://www.medical.philips.com.

[7] (2005). Worldheart association. http://www.worldheart.org.

[8] (2005). Yale atlas of echo. http://info.med.yale.edu/intmed/cardio/echo_atlas/views/.

[9] Abd-Elmoniem, K., Youssef, A.-B., and Kadah, Y. (2002). Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion. *IEEE Transactions on Biomedical Engineering*, 49(9):997–1014.

[10] Abed-Meraim, K., Cardoso, J. F., Gorokhov, A., Loubaton, P., and Moulines, E. (1997). On subspace methods for blind identification of single-input multiple-output fir systems. *IEEE Transactions on Signal Processing*, 45(1):42 – 55.

[11] Achim, A., Bezerianos, A., and Tsakalides, P. (2001). Novel bayesian multiscale method for speckle removal in medical ultrasound images. *IEEE Transactions on Medical Imaging*, 20(8):772–783.

[12] Adam, D. and Michailovich, O. (2002). Blind deconvolution of ultrasound sequences using nonparametric local polynomial estimates of the pulse. *IEEE Transactions on Biomedical Engineering*, 49(2):118–131.

[13] Anderson, M. E. and Trahey, G. E. (2000). A seminar on k-space applied to medical ultrasound. Technical report, Department of Biomedical Engineering Duke University.

[14] Andrews, H. C. and Hunt, B. R. (1977). *Digital Image Restoration*. Prentice Hall, New Jersey.

[15] Angelini, E. D., Laine, A., Takuma, S., Holmes, J., and Homma, S. (2001). Lv volume quantification via spatio-temporal analysis of real-time 3d echocardiography. *IEEE Transactions on Medical Imaging*, 20:457–469.

[16] Arfken, G. (1985). *Mathematical Methods for Physicists, 3rd ed.* Academic Press, Orlando, FL.

[17] Ayers, G. and Dainty, J. (1988). Iterative blind deconvolution method and its applications. *Optics Letters*, 13(7):547–549.

[18] Barrett, R., Berry, M., Chan, T. F., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., and der Vorst, H. V. (1994). *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition.* SIAM, Philadelphia, PA.

[19] Bates, R. H. T. (1982). Astronomical speckle imaging. *Phys Rep.*, 90(4):203 – 297.

[20] Bayes, T. (1764). An essay toward solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53:370–418.

[21] Bellini, S. (1994). *Blind Deconvolution*, chapter Bussgang techniques for blind deconvolution and restoration. Prentice-Hall.

[22] Bercoff, J., Tanter, M., and Fink, M. (2004). Supersonic shear imaging: a new technique for soft tissue elasticity mapping. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 51(5):606–615.

[23] Berenstein, C. and Patrick, E. (1990). Exact deconvolution or multiple operators – an overview plus performance characterizations for imaging sensors. *IEEE Proc. in Multidimensional Signal Processing*, 78(4):723–734.

[24] Besl, P. J. and McKay, N. D. (1992). A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256.

[25] Black, M. and Rangarajan, A. (1996). On the unification of line process, outlier rejection, and robust statistics with applications in early vision. *Internat. J. Comput. Vision*, 19(1):57–91.

[26] Black, M. J., Sapiro, G., Marimont, D. H., and Heeger, D. (1998). Robust anisotropic diffusion. *IEEE Transactions on Image Processing*, 7(3):421–432.

[27] Blake, A. and Zisserman, A. (1987). *Visual Reconstruction.* MIT Press, Cambridge, MA.

[28] Blanc-Féraud, L. (2000). *Sur quelques problèmes inverses en traitement d'image.* Habilitation à diriger des recherches, Université de Nice-Sophia Antipolis.

[29] Blanc-Féraud, L., Charbonnier, P., Aubert, G., and Barlaud, M. (1995). Nonlinear image processing: modeling and fast algorithm for regularization with edge detection. In *Proceedings of the ICIP 1995*, pages 474–477.

[30] Bloch, I. (1996). Information Combination Operators for Data Fusion: A Comparative Review wth Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 26(1):52–57.

[31] Bookstein, F. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585.

[32] Burckhardt, C. B. (1978). Speckle in ultrasound b-mode scans. *IEEE Transactions on Sonics and Ultrasonics*, 25(1):1–6.

[33] Burt, P. (1992). A gradient pyramid basis for pattern selective image fusion. In *Proc. of the Society for Information Display conference*, volume 16, pages 467–470.

[34] Burt, P. and Adelson, A. (1983). The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31:532–540.

[35] Carmo, B. S., Prager, R. W., Gee, A. H., and Berman, L. H. (2002). Speckle detection for 3d ultrasound. *Ultrasonics*, 40(1-8):129–132.

[36] Chakravarti, Laha, and Roy (1967). *Handbook of Methods of Applied Statistics*, volume 1. John Wiley and Sons.

[37] Chang, M., Tekalp, A. M., and Erdem, A. (1991). Blur identification using the bispectrum. *IEEE Transactions on Signal Processing*, 39(10):2323 – 2325.

[38] Charbonnier, P., Blanc-Féraud, L., Aubert, G., and Barlaud, M. (1997). Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions in Image Processing*, 6(2):298–311.

[39] Chen, C. H. and Qidwai, U. (2002). Recent trends in 2d blind deconvolution for nondestructive evaluation. volume 2, pages 989 – 992.

[40] Cincotti, G., Loi, G., and Pappalardo, M. (2001). Frequency decomposition and compounding of ultrasound medical images with wavelet packets. *IEEE Transactions on Medical Imaging*, 20(8):764–771.

[41] Conchello, J. A. (1998). Super-resolution and convergence properties of the expectation-maximization algorithm for maximum-likelihood deconvolution of incoherent images. *Journal of the Optical Society of America*, 15(10):2609–2620.

[42] Criton, A., Steel, R., Hoskins, P. R., McDicken, W. N., and Routh, H. F. (2002). Real time vector doppler for tissue motion. In *Proceedings of IEEE Ultrasonics Symposium*, volume 2, pages 1529–1534.

[43] Donoho, D. L. (1995). De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627.

[44] Duan, Q., Angelini, E. D., and Laine, A. (2004). Assessment of fast anisotropic diffusion and scan conversion of real-time three-dimensional spherical ultrasound data for visual quality and spatial accuracy. In *SPIE International Symposium on Medical Imaging*, volume 5370, pages 526–537.

[45] Dutt, V. and Greenleaf, J. (1996). Adaptive speckle reduction filter for log-compressed b-scan images. *IEEE Transactions on Medical Imaging*, 15(6):802 – 813.

[46] Elad, M. (2002). On the bilateral filter and ways to improve it. *IEEE Transactions on Image Processing*, 11(10):1141–1151.

[47] Evans, A. and Nixon, M. S. (1993). Temporal methods for ultrasound speckle reduction. In *IEE Seminar on Texture analysis in radar and sonar*, volume 1, pages 1–6.

[48] Evans, D. H. and McDicken, W. N. (2000). *Doppler Ultrasound: Physics, Instrumental and Signal Processing (2nd ed.)*. John Wiley and Sons, West Sussex, England.

[49] Farsiu, S., Robinsonand, M. D., Elad, M., and Milanfar, P. (2004). Fast and robust multiframe super resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344.

[50] Ferlay, J. (2001). Globocan 2000. In *IARC CancerBase*, number 5. World Health Organization, International Agency for Research on Cancer.

[51] Forsythe, G. E., Malcolm, M. A., and Moler, C. B. (1976). *Computer Methods for Mathematical Computations*. Prentice-Hall.

[52] Frost, V. S., Stiles, J. A., Shanmugan, K. S., and Holtzman, J. C. (1982). A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Trans. Pattern Anal. Machine Intell.*, 4(2):157–165.

[53] Galatsanos, N. P., Mesarovic, V. Z., and Kang, M. G. (1994). *On Image Restoration Techniques for Medical Imaging*, volume 3, chapter 1, pages 1–68. Gordon and Breach Publishers.

[54] Garra, B., Cespedes, E. I., Ophir, J., Spratt, S. R., Zuurbier, R. A., Magnant, C. M., and Pennanen, M. F. (1997). Elastography of breast lesions: Intial clinical results. *Radiology*, 202:79–86.

[55] Geman, S. and Geman, D. (1984). Sthocastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741.

[56] Giannakis, G. B. and Robert, H. W. J. (2000). Blind identification of multichannel fir blurs and perfect image restoration. *IEEE Transactions on Image Processing*, 9(11):1877–1896.

[57] Gonzalez, R. C. and Woods, R. E. (2002). *Digital Image Processing (2nd ed.)*. Prentice-Hall.

[58] Goodman, J. W. (1968). *Introduction to Fourier Optics*. McGraw-Hill, San Francisco.

[59] Grau, V. and Noble, J. A. (2005). Adaptive multiscale ultrasound compounding using phase information. In *Proceedings of MICCAI 2005*, pages 589–596.

[60] Guo, H., Odegard, J., Lang, M., Gopinath, R., Selesnick, I., and Burrus, C. (1994). Wavelet based speckle reduction with application to sar based atd/r. In *IEEE International Conference Image Processing*, volume 1, pages 75 – 79.

[61] Gupta, S., Chauhan, R. C., and Sexana, S. C. (2001). Wavelet-based statistical approach for speckle reduction in medical ultrasound images. *Med. Biol. Eng. Comput.*, 42:189–192.

[62] Gurelli, I. and Chrysostomos, N. L. (1995). Evam: an eigenvector-based algorithm for multichannel blind deconvolution of input colored signals. *IEEE Transactions on Signal Processing*, 43(1):134–149.

[63] Hadamard, J. (1902). Sur les problèmes aux derivées partielles et leur signification physique. *Bull. Univ. of Princeton*, pages 49–52.

[64] Hao, X., Gao, S., and Gao, X. (1999). A novel multiscale nonlinear thresholding method for ultrasonic speckle suppressing. *IEEE Transactions on Medical Imaging*, 18(9):787–794.

[65] Harikumar, G. and Bresler, Y. (1999). Perfect blind restoration of images blurred by multiple filters: theory and efficient algorithms. *IEEE Transactions on Image Processing*, 8(2):202–219.

[66] Haykin, S., editor (1994). *Blind Deconvolution*. Prentice-Hall.

[67] Heath, R. W., Halford, S. D., and Giannakis, G. B. (1996). Adaptive blind channel identification of fir channels for viterbi decoding. In *Proc. 30th. Annu. Asilomar Conf. Signals, Systems, Computers*, pages 3–6.

[68] Högborn, J. (1974). aperture synthesis with a non-regular distribution of interferometer baselines. *Astrophys. J. Suppl. Ser.*, 15:417–426.

[69] Hieken, T. J., Harrison, J., Herreros, J., and Velasco, J. M. (2001). Correlating sonography, mammography, and pathology in the assessment of breast cancer size. *The American Journal of Surgery*, 182(4):351–354.

[70] Holmes, T. J. (1992). Blind deconvolution of speckle images quantum-limited incoherent imagery: maximum-likehood approach. *J. Opt. Soc. Am.*, 9:1052 – 1061.

[71] Idier, J. and Goussard, Y. (1993). Multichannel seismic deconvolution. *IEEE Trans. Geoscience Remote Sensing*, 31(5):961–979.

[72] Insana, M. F., Wagner, R. F., Garra, B. S., Brown, D. G., and Shawker, T. H. (1986). Analysis of ultrasound image texture via generalized rician statistics. *Optical Engineering*, 25(6):743–748.

[73] Issacson, E. and Keller, H. (1966). *Analysis of Numerical Methods*. Wiley, New York, NY.

[74] Izzetoglu, M., Onaral, B., and Bilgütay, N. (2000). Wavelet domain least squares deconvolution for ultrasonic backscattered signals. In *Proceedings of the 22th Annual EMBS International Conference*, pages 321–477, Chicago, USA.

[75] Jain, A. K. (1989). *Fundamentals of Digital Image Processing*. Prentice Hall, NJ.

[76] Jansen, C., Arigovindan, M., Sühling, M., Marsch, S., Unser, M., and Hunziker, P. (2003). Multidimensional, multistage wavelet footprints: A new tool for image segmentation and feature extraction in medical ultrasound. In Sonka, M. and Fitzpatrick, J., editors, *Progress in Biomedical Optics and Imaging, vol. 4, no. 23*, volume 5032 of *Proceedings of the SPIE International Symposium on Medical Imaging: Image Processing (MI'03)*, pages 762–767, San Diego CA, USA. Part II.

[77] Jensen, J. (1996a). *Estimation of Blood Velocities Using Ultrasound: A Signal Processing approach*. Cambridge University Press.

[78] Jensen, J. (1996b). Field: A program for simulating ultrasound systems. In *10th Nordic-Baltic Conference on Biomedical Imaging*, volume 34, pages 351–353.

[79] Kao, C., Pan, X., Hiller, E., and Chen, C. (1998). A bayesian approach for edge detection in medical ultrasound images. *IEEE Transactions on Nuclear Science*, 45(6):3089–3096.

[80] Kaplan, D. and Ma, Q. (1994). On the statistical characteristics of the log-compressed rayleigh signals: Theorical formulation and experimental results. *J. Acoust. Soc. Amer.*, 95:1396 – 1400.

[81] Katsaggelos, A. K. (1989). Iterative image restoration algorithms. *Optical Engineering*, 27(7):735–748.

[82] Katsaggelos, A. K. (1990). A multiple input image restoration approach. *J. Visual Commun. Image Represent*, 1:93–103.

[83] Katsaggelos, A. K. (1991). *Digital Image Restoration*. Springer-Verlag, New York.

[84] Kautsky, J., Flusser, J., Zitová, B., and Simberová, S. (2002). A new waveletbased measure of image focus. *Pattern Recognit. Lett.*, 23:1785–1794.

[85] Kirkhorn, J., Frigstad, S., and Torp, H. (2000). Comparison of pulse inversion and second harmonic for ultrasound contrast imaging. In *Proceedings of the IEEE Ultrasonics Symposium*, volume 2, pages 1897–1901.

[86] Künsch, H. R. (1994). Robust priors for smoothing and image restoration. *Ann. Inst. Statist. Math.*, 46(1):1–19.

[87] Kovesi, P. (1999). Phase preserving denoising of images. In *Proc. of the Australian Pattern Recognition Society Conference: DICTA'99*, pages 212–217, Perth, WA.

[88] Krissian, K., Vosburgh, K., Kikinis, R., and Westin, C.-F. (2004). Anisotropic diffusion of ultrasound constrained by speckle noise model. Technical Report 0004, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Laboratory of Mathematics in Imaging. ISSN.

[89] Krissian, K., Vosburgh, K., Kikinis, R., and Westin, C.-F. (2005). Speckle-constrained anisotropic diffusion for ultrasound images. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 547–552, San Diego CA, USA.

[90] Krücker, J. F., LeCarpentier, G. L., Fowlkes, J. B., and Carson, P. L. (2002). Rapid elastic image registration for 3-d ultrasound. *IEEE Transactions on Medical Imaging*, 21:1384 – 1394.

[91] Kundur, D. and Hatzinakos, D. (1996). Blind image deconvolution. *IEEE Transactions on Signal Processing*, 13(3):43 – 64.

[92] Kundur, D. and Hatzinakos, D. (1998). A novel blind deconvolution scheme for image restoration using recursive filtering. *IEEE Transactions on Signal Processing*, 46(2):375–390.

[93] Lagendijk, R. L., Tekalp, A. M., and Biemond, J. (1990). Maximum likelihood image and blur identification: a unifying approach. *Optical Engineering*, 29(5):422–435.

[94] Lane, R. G. (1992). Blind deconvolution of speckle images. *J. Opt. Soc. Am.*, 9:1508 – 1514.

[95] Lane, R. G. and Bates, R. H. T. (1987). Automatic multidimensional deconvolution. *J. Opt. Soc. Am.*, 4(1):180 – 188.

[96] LeCarpentier, G. L., Tridandapani, P. B., Fowlkes, J. B., Roubidoux, M. A., Moskalik, A. P., and Carson, P. L. (1999). Utility of 3-d ultrasound in the discrimination and detection of breast cancer. *Radiological Society of North America Electronic Journal*, 3.

[97] Lee, J. S. (1980). Digital image enhancement and noise filtering by using local statistics. *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI(2):165–168.

[98] Lee, S., Wolberg, G., and Shin, S. Y. (1999). Scattered data interpolation with multilevel b-splines. *IEEE Trans. Visualization Comput. Graph.*, 3:228–244.

[99] Legget, M. E. and Leotta, D. F. (1998). System for quantitative 3d echocardiography of the left ventricle based on a magnetic-field position and orientation sensing system. *IEEE Trans. Biomed. Eng.*, 45(4):495–504.

[100] Leotta, D. F. and Martin, R. W. (2000). Three-dimensional spatial compounding of ultrasound scans with incidence angle weighting. *Ultrasonic Imaging*, 22:1–19.

[101] Li, H., Manjunath, B. S., and Mitra, S. K. (1994). Multi-sensor image fusion using the wavelet transform. In *Proc. of the IEEE International Conference Image Processing*, volume 1, pages 51–55.

[102] Loupas, T., Pye, S. D., and McDicken, W. N. (1989). Deconvolution in medical ultrasonics: practical considerations. *Phys. Med. Biol.*, 34:1691 – 1700.

[103] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P. (1997). Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16:187–198.

[104] Makela, T., Clarysse, P., Sipila, O., Pauna, N., Pham, Q. C., Katila, T., and Magnin, I. (2002). Review of cardiac image registration methods. *IEEE Transactions on Medical Imaging*, 21(9):1011 – 1021.

[105] Mallat, S. (1997). *A Wavelet Tour of Signal Processing*. AP Professional, London.

[106] Martinoli, C., Pretolesi, F., Crespi, G., Bianchi, S., Gandolfo, N., Valle, M., and Derchi, L. (1998). Power doppler sonography: clinical applications. *Eur. J. Radiol.*, 27(2):133–140.

[107] McCallum, B. C. (1990). Blind deconvolution by simulated annealing. *Optics Communications*, 75(2):101–105.

[108] Meinel, E. (1986). Origins of linear and nonlinear recursive restoration algortihms. *Journal Optical Soc. Amer.*, 3:787 – 799.

[109] Michailovich, O. and Adam, D. (2003). Robust estimation of ultrasound pulses using outlier-resistant de-noising. *IEEE Transactions on Medical Imaging*, 22(3):368–381.

[110] Montagnat, J., Sermesant, M., Delingette, H., Malandain, G., and Ayache, N. (2003). Anisotropic filtering for model-based segmentation of 4d cylindrical echocardiographic images. *Pattern Recognition Letters*, 24(4-5):815–828.

[111] Moore, S. (2001). Better breast cancer detection. *Spectrum, IEEE*, 38(5):50–54.

[112] Moulines, E., Duhamel, P., Cardoso, J. F., and Mayrargue, S. (1995). Subspace methods for the blind identification of multichannel fir filters. *IEEE Transactions on Signal Processing*, 43(2):516–525.

[113] Mrázek, P., Weickert, J., and Steidl, G. (2003). Correspondences between wavelet shrinkage and nonlinear diffusion. In Griffin, L. D. and Lillholm, M., editors, *Scale-Space 2003. Lecture Notes in Computer Science*, volume 2695, pages 101–116. Springer.

[114] Mulet-Parada, M. and Noble, J. A. (2000). 2D + T Acoustic Boundary Detection in Echocardiography. *Medical Image Analysis*, 4(1):21–30.

[115] Nikolova, M. (2000). Thresholding implied by truncate quadratic regularization. *IEEE Transactions on Signal Processing*, 48(12):3437–3450.

[116] O'Donnell, M. and Silverstein, S. D. (1988). Optimum displacement for compound image generation in medical ultrasound. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 35:470–476.

[117] Ophir, J., Alam, S., Garra, B., Kallel, F., Konofagou, E., Krouskop, T., Merritt, C., Righetti, R., Souchon, R., Srinivasan, S., and Varghese, T. (2002). Elastography : Imaging the elastic properties of soft tissues with ultrasound. *J. Med. Ultrasonics*, 29:155–171.

[118] Ophir, J., Cespedes, I., Ponnekanti, H., Yazdi, Y., and Li, X. (1991). Elastography : a quantitative method for imaging the elasticity of biological tissues. *Ultrasonic imaging*, 13:111–134.

[119] Pai, H. and Bovik, A. C. (2001). On eigenstructure-based direct multichannel blind image restoration. *IEEE Transactions on Image Processing*, 10(10):1434–1446.

[120] Pai, H., Havlicek, J., and Bovik, A. (1998). Generically sufficient conditions for exact multichannel blind image restoration. In *Proc. ICASSP 98*.

[121] Panci, G., Campisi, P., Colonnese, S., and Scarano, G. (2003). Multichannel blind image deconvolution using the bussgang algorithm: spatial and multiresolution approaches. *IEEE Transactions on Image Processing*, 12(11):1324– 1337.

[122] Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–638.

[123] Piella, G. (2003). A general framework for multiresolution image fusion: from pixels to regions. *Information Fusion*, 9:259–280.

[124] Prager, R. W., Gee, A. H., Treece, G. M., , and Berman, L. H. (2002). Analysis of speckle in ultrasound images using fractional order statistics of the homodyned k-distribution. *Ultrasonics*, 40(1-8):133–137.

[125] Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1992a). *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge (UK) and New York, 2nd edition.

[126] Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1992b). *Robust Estimation*, chapter 15.7, pages 694–700. Cambridge University Press, Cambridge, England, second edition.

[127] Reeves, S. and Mersereau, R. (1992). Blur identification by the method of generalized cross-validation. *IEEE Transactions on Image Processing*, 1:301–311.

[128] Roche, A., Pennec, X., Malandain, G., and Ayache, N. (2001). Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE Transactions on Medical Imaging*, 20(10):1038–1049.

[129] Rohling, R., Gee, A., and Berman, L. (1996). Three-dimensional spatial compounding of ultrasound images. *Medical Image Analysis*, 1(3):177–193.

[130] Rudin, L., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268.

[131] Rueckert, D., Somoda, I., Hayes, C., Hill, D., Leach, M., and Hawkes, D. (1999). Nonrigid registration using free-form deformations: Applications to breast mr images. *IEEE Transactions on Medical Imaging*, 18(8):712–721.

[132] Sapia, M., Loew, L., Fox, M., and Schaff, J. (1998). Ultrasound deconvolution using adaptive inverse filtering. In *Proceedings of the 12th IEEE Symposium on Computer-Based Medical Systems*, pages 248–253.

[133] Sarrut, D. (2000). *Recalage multimodal et plate-forme d'imagerie médicale à accès distant*. PhD thesis, Université Lumiére Lyon 2, France.

[134] Shekhar, R., Zagrodsky, V., Garcia, M. J., and Thomas, J. D. (2004). Registration of real-time 3-d ultrasound images of the heart for novel 3-d stress echocardiography. *IEEE Trans. Med. Imaging*, 23(9):1141–1149.

[135] Shepp, L. A. and Vardi, Y. (1982). Maximum likelihood reconstruction for emission tomography. *IEEE Transactions On Medical Imaging*, 1:113–121.

[136] Smith, S. W., Wagner, R., Sandrik, J. M., and Lopez, H. (1983). Low contrast detectability and contrast/detail analysis in medical ultrasound. *IEEE Transactions on Sonics and Ultrasonics*, 30(3):164–173.

[137] Sohn, C., Blohmer, J., and Hamper, U. (1999). *Breast Ultrasound: A Systematic Approach to Technique and Image Interpretation*. Thieme Medical Publishers, Stuttgart, Germany.

[138] Soler, P., Delso, G., Villain, N., Angelini, E., and Bloch, I. (2006). Superresolution Spatial Compounding Techniques, with Application to 3D Breast Ultrasound Imaging. In *Proceedings of SPIE 2006 (to appear)*, San Diego, CA.

[139] Soler, P., Gerard, O., Allain, P., Saloux, E., Angelini, E., and Bloch, I. (2005a). Comparison of fusion techniques for 3D+T echocardiography acquisition from different acoustic windows. In *Proceedings of Computers In Cardiology (to appear)*, Lyon, France.

[140] Soler, P., Villain, N., Bloch, I., and Angelini, E. (2005b). Volume reconstruction of breast echography from anisotropically degradated scans. In *Proceedings of the IASTED International Conference on Biomedical Engineering*, volume 9, pages 349–355, Innsbruck, Austria.

[141] Sroubek, F. and Flusser, J. (2003). Multichannel blind iterative image restoration. *IEEE Transactions on Image Processing*, 12(9):1094–1106.

[142] Sroubek, F. and Flusser, J. (2005). Multichannel blind deconvolution of spatially misaligned images. *IEEE Transactions on Image Processing*, 14(7):874–883.

[143] Starck, J. L. and Murtagh, F. (1994). Image restoration with noise suppression using the wavelet transform. *Astronomy and Astrophysics*, 288:343–348.

[144] Starck, J. L. and Pantin, E. (2002). Deconvolution in astronomy: A review. *Publications of the Astronomical Society of the Pacific*, 114:1051–1069.

[145] Steidl, G., Weickert, J., Brox, T., Mrázek, P., and Welk, M. (2004). On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and sides. *SIAM Journal on Numerical Analysis*, 42(2):686–713.

[146] Subbarao, M. and Choi, T. (1995). Accurate recovery of threedimensional shape from image focus. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17:266–274.

[147] Subbarao, M. and Tyan, J. (1998). Selecting the optimal focus measure for autofocusing and depth-from-focus. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20:864–870.

[148] Tauber, C., Batatia, H., and Ayache, A. (2004). A robust speckle reducing anisotropic diffusion. *International Conference on Image Processing ICIP '04*, 1(1):247–250.

[149] Taxt, T. (2001). Three-dimensional blind deconvolution of ultrasound images. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 48(4):867–871.

[150] Taxt, T. and Strand, J. (2001). Two-dimensional noise-robust blind deconvolution of ultrasound images. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 48(4):861–866.

[151] Tikhonov, A. N. and Arsenin, V. Y. (1977). *Solution of Ill-posed problems*. W. H. Winston and Sons, Washington, D.C.

[152] Tomasi, C. and Manduchi, R. (2003). Bilateral filtering for gray and color images. In *Proceedings of IEEE International Conference on Computer Vision*, pages 59–66, Bombay, India.

[153] Tong, L. and Perreau, S. (1998). Multichannel blind identification: From subspace to maximum likelihood methods. *Proceedings of the IEEE*, 86(10):1951–1968.

[154] Tong, L., Xu, G., and Kailath, T. (1994). Blind identification and equalization based on second order statistics: A time domain approach. *IEEE Trans. Inf. Theory*, 40:340–349.

[155] Trahey, G. E., Smith, S. W., and von Ramm, . T. (1986). Speckle reduction in medical ultrasound via spatial compounding. In *Proc. 14th SPIE on Medical Applications*, pages 626–637.

[156] Veldhuizen, T. L. and Jernigan, M. E. (1998). Grid filters for local nonlinear image restoration. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP'98)*.

[157] Vilkomerson, D., Greenleaf, J., and Dutt, V. (1995). Towards a resolution metric for medical ultrasonic imaging. volume 2, pages 1405 – 1410, Chicago, USA.

[158] Villain, N., Goussard, Y., Idier, J., and Allain, M. (2003). Three-dimensional edge-preserving image enhancement for computed tomography. *IEEE Trans. Medical Imaging*, 22(10):1275–1287.

[159] Wagner, R., Smith, S. W., Sandrik, J. M., and Lopez, H. (1983). Statistics of speckle in ultrasound b-scans. *IEEE Transactions on Sonics and Ultrasonics*, 30(3):156–163.

[160] Wan, S., Raju, B. I., and Srinivasan, M. A. (2003). Robust deconvolution of high-frequency ultrasound images using higher-order spectral analysis and wavelets. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 50(10):1286– 1295.

[161] Weickert, J. (1995). *Computer Analysis of Images and Patterns*, volume 970 of *Lecture Notes in Computer Science*, chapter Mutiscale texture enhancement, pages 230–237. Springer, Berlin.

[162] Wilhjelm, J. E., Jensen, M. S., Brandt, T., Sahl, B., Martinsen, K., Jespersen, S. K., and Falk, E. (2000). Some imaging strategies in multi-angle spatial compounding. In *Proc. of the IEEE International Ultrasonics Symposium*, volume 2, pages 1615–1618.

[163] Wirawan, I., Abed-Meraim, K., Maitre, H., and Duhamel, P. (2002). Blind multichannel image restoration using oblique projections. In *Proceedings of the Sensor Array and Multichannel Signal Processing Workshop*, pages 125–129.

[164] Xiao, G., Brady, J. M., Noble, J. A., Brucher, M., and English, R. (2002). Non-rigid registration of 3-D free-hand ultrasound images of the breast. *IEEE Transactions on Medical Imaging*, 21(4):405–412.

[165] Ye, X., Noble, J. A., and Atkinson, D. (2002). 3D freehand echocardiography for automatic left ventricle reconstruction and analysis based on multiple acoustic windows. *IEEE Trans. Med. Imaging*, 21(9):1051–1058.

[166] Yu, Y. and Acton, S. T. (2002). Speckle reducing anisotropic diffusion. *IEEE Transactions on Image Processing*, 11(11):1260–1270.

[167] Zagzebski, J. A. (1996). *Essentials of Ultrasound Physics*. Mosby.

Citations index by article

List of Figures

List of Tables