بِسْمِ اللهِ الرَّحْمٰنِ الرَّحِيْمِ

Islamic University of Technology (IUT)
Department of Computer Science and Engineering (CSE)

Strain Estimation and Detection of Cancerous Breast Lesion through ultrasound image

Authors
Abrar Faiyaz - 124405
&
Md. Samiul Bashar Anan - 124451

Supervisor
Dr. Md. Hasanul Kabir
Associate Professor, Department of CSE
Islamic University of Technology (IUT)

A thesis submitted to the Department of CSE
in partial fulfillment of the requiremnts for the degree of B.Sc. Engineering in CSE
Academic Year: 2015-16

November - 2016

Certification

This is to certify that the work presented in this thesis is the outcome of the analysis and experiments carried out by Abrar Faiyaz and Md. Samiul Bashar Anan under the supervision of Dr. Md. Hasanul Kabir, Associate Professor, Department of Computer Science and Engineering, Islamic University of Technology (IUT), Dhaka, Bangladesh. It is also declared that neither of this thesis nor any part of this thesis has been submitted anywhere else for any degree or diploma. Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given.

Authors:

Abrar Faiyaz
Student ID - 124405

Md. Samiul Bashar Anan
Student ID - 124451

Supervisor:

Dr. Md. Hasanul Kabir
Associate Professor, Department of Computer Science and Engineering
Islamic University of Technology (IUT)

Abstract

Currently in the domain of breast cancer detection, biopsy and ultrasound imaging technique is being used. Between these two techniques biopsy is costly, time consuming and tissue damaging for patients. Ultrasound imaging technique is an easy alternative which needs experts to accurately detect breast lesion and it's type. In the BIRAD method of identification, Sonologists look for certain features of the malignant lesions. These malignant lesions in static image has some significant features which with 15-30% cases result in contradictory results. Most important feature that a breast lesion posses is it's tissue property ie. tissue destiny and the deforming mechanism. These properties are associated with the strain value when pressure is applied to the tissue. There are different ways the strain value can be calculated by using multiple images with different strain values. Generally,the malignant lesions exhibits more rigidity than the benign ones. From this motivation, optical flow algorithm has been implemented to calculate the strain value and compare it's measure of mean related to accuracy with cross correlation method used to calculate the strain values. From this motivation, we have figured out strain with higher accuracy adopting our "M-Filtered B-Mode Image Tissue Elasticity" estimation method. Higher the accuracy, higher will be the mean of the calculated strain. Satisfactory SNR (dB) results in our implemented method of optical flow which is promising for a classification with higher accuracy.

Contents

List of Figures

List of Tables

Chapter 1

1 Introduction

In this chapter, we first present an overview of our thesis that includes the signification of the problem and the problem statement in detail. Research challenges to be faced in the whole scenario is also discussed based on the problem statement. Thesis objectives, motivations and our contribution are noted in sections. The end of this chapter has the description of the organization of the thesis.

1.1 Overview

1.1.1 Ultrasound Imaging

Medical ultrasound (also known as diagnostic sonography or ultrasonography) is a diagnostic imaging technique based on the application of ultrasound. It is used to see internal body structures such as tendons, muscles, joints, vessels and internal organs. Ultrasound is now using in different sectors of medical imaging such as- Gall bladder Stone (Polyp/Worm) Detection, Kidney Stone/Tumor/Parenchymal disease Detection, Urinary bladder Stone/Tumor Detection, Breast and Ovarian tumor Detection, Blood vessel stenosis/obstruction Detection.

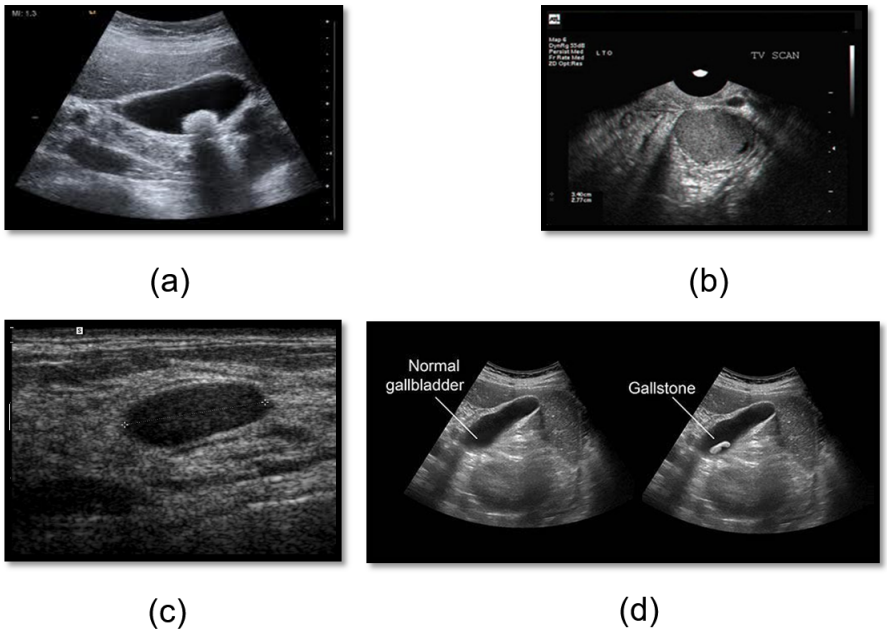


Figure 1.1: (a) Kidney Stone (b) Ovarian Tumor (c) Breast Tumor (d) Gall Bladder Stone

1.1.2 Ultrasound Image Generation

Ultrasound probe sends high frequency sound wave through tissues. There is a receiver called piezoelectric crystal which receive different types of sound echo. They are-

- Hyper echoic: Much brighter appearance (bone,air,capsules)
- Iso echoic: Structures bearing similar echogenicities(muscle,fat)
- Hypo echoic: Much darker appearance(fluid)

Figure 1.2: Image generation in ultrasonography

1.1.3 Breast Cancer Lesion : Properties

The cancer affected region that forms a lump is called a cancerous Lesion. Breast cancer is two types-

1. Benign(Early stage) lump:
   - Well circumscribed, hyperechoic tissue: 100%
   - Wider than deep: 99%
   - Gently curving smooth lobulations (lesser than 3 in a wider than deep nodule, i.e. D/W ratio lesser than1): 99%
   - thin echogenic pseudocapsule in a wider than deep nodule: 99%
2. Malignant(Late stage):
   - Markedly hypoechoic nodule: 70%
   - Sonographic posterior acoustic shadowing: 50%
   - Branching pattern: 30%
   - Punctate calcifications: 25%
   - Duct extension: 25%
   - Heterogeneous echotexture
   - Compressibility: in general terms, benign lesions compress with transducer pressure and malignant lesions displace the breast tissue without changing in height; this is the basis for elastography

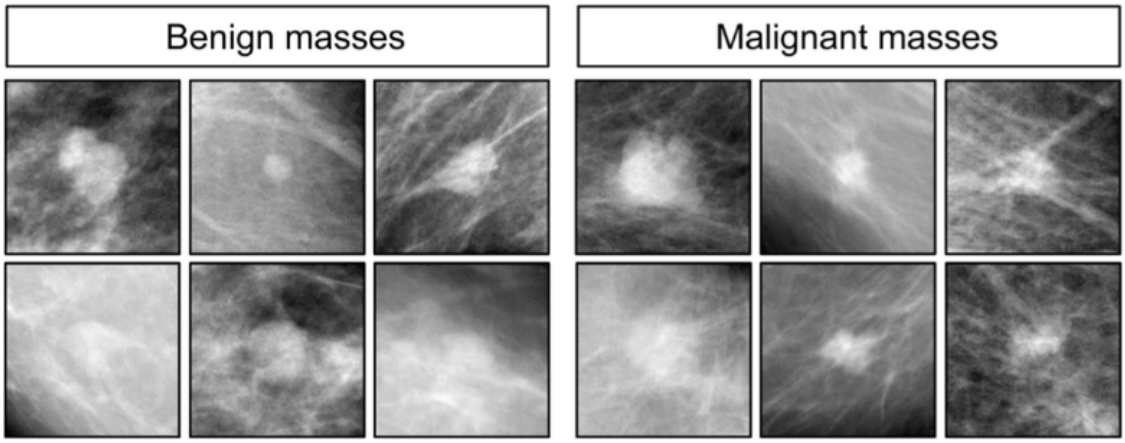


Figure 1.3: Benign and Malignant masses

Strain estimation is the measure of deformation that the tissue undergoes when pressure is applied on it. Properties of tissue materials and idea regarding their densities and rigidity can be known through the means of strain estimation. Ultrasound images are now becoming popular in the radio frequency domain as well as in the image domain for the purpose of research. Basically, the method is cheaper with respect to the alternatives and hustle-free. Its easy means of operation makes it viable to work with. Strain estimation and detection of cancerous breast lesion through ultrasound image requires the following steps - For detection the steps necessary are preprocessing, feature extraction, classification and post-processing as shown in figure 1.1. Besides these steps, a key issue in successful detection of cancerous breast lesion is to find an effective feature extraction method, which will provide a robust and efficient means for the classification of malignant and benign cancerous lesions for the classifier to train and test. Being constrained with our dataset we focused on the performance of the feature that we need to extract based on tissue properties i.e. strain estimation.

The strain estimation feature can be extracted by means of video image or the images taken sequentially based on the percentage of the strain. When the pressure applied is known and duration of the impulse is known, the feature can be estimated using the video data. On the contrary, if any of the mentioned parameters are unknown then static multiple images whose compression rates are known can be used to extract the strain value. Here, the duration of the impulse was unknown and we worked with multiple static images where the compression rate was given.

Strain estimation technique is based on the 2D strain values. Lateral and Axial strain values are respectively along Y and X axis.

The elasticity of a material describes its tendency to resume its original size and shape after being subjected to a deforming force or stress. Fluids resist a change in volume, but not in shape: they possess only volume elasticity. Solids resist changes in shape and volume, they possess rigidity or shear elasticity, as well as volume elasticity. The change in size or shape is known as the strain, which is expressed as a ratio (e.g. the change in length per unit length). The strain is produced by a system of forces; the force acting on unit area is known as the stress.

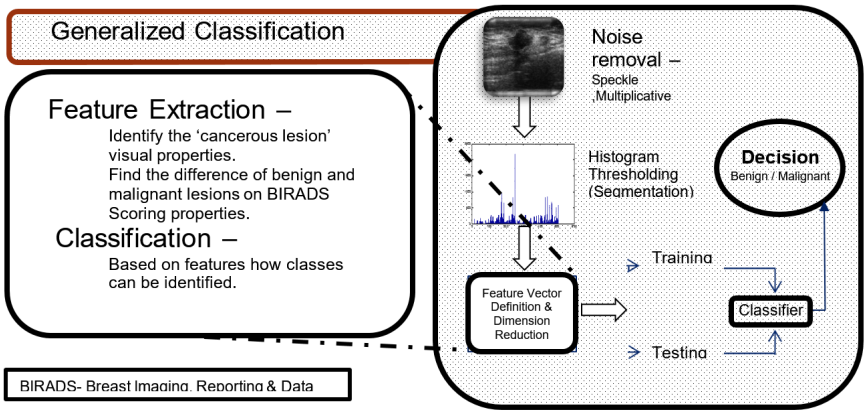


Figure 1.4: Components of a generic means of classification and focusing our area of feature extraction and detection

1.2 Significance of the problem

Two techniques are used in detecting cancerous breast lesion.

1. Biopsy surgery
2. Ultrasound Imaging[16]

A biopsy is a small operation done to remove tissue from an area of concern in the body. If your doctor feels anything suspicious in your breast, or sees something suspicious on an imaging study, he or she will order a biopsy. The tissue sample is examined by a pathologist (a doctor who specializes in diagnosing disease) to see whether or not cancer cells are present. If cancer is present, the pathologist can then look at the cancer's characteristics. The biopsy will result in a report that lays out all of the pathologist's findings.

Imaging studies such as mammography and ultrasound, often along with physical exams of the breast, can lead doctors to suspect that a person has breast cancer. In between the two imaging techniques ultrasound is safer, cheaper, more convenient and faster than mammography. In case of Biopsy, statistics says that only 30% of the patients are found having cancerous lesion(malignant).

Biopsy surgery is costly, and for developing and underdeveloped country this facility is not available everywhere. On the other hand, Ultrasound is very cheap and available everywhere. If the results are improved in ultrasound breast cancer detection, then the rate of biopsy can be reduced. This will minimizes the cost and increase the availability. So our research is focused on generating significant features that will differentiate between the benign and malignant lesion. In this case, the strain feature gives distinct behaviour in those two kind of lesions. So by estimating the strain from ultrasound breast images by applying pressure, we can use the strain to classify the cancerous lesion.

1.3 Research Challenges

There are different types of ultrasound devices[16] which use different frequencies. So that the RF data varies in different devices. It is difficult to generalize the method for all kind of frequencies. Another thing is many features can be extracted from the image generated from ultrasound RF data. All features are not relevant. All of them can not distinctly classify between the two cancer types[16]. So reducing the feature dimension gives better result. Finding out the relevant features is

an important issue. Again in the ultrasound breast image there are fatty tissues having the similar texture like the lesion. It is difficult to find out the lesion region because the ultrasound image is very noisy so that it contains less information to differentiate between those fatty tissues and lesion. There is also shadowing effect[14] available around the lesion region. It creates difficulties to segment out the proper region of interest.

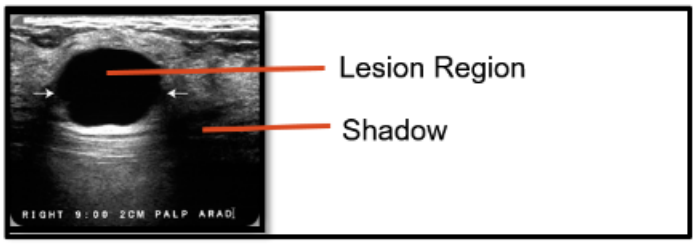


Figure 1.5: Shadowing problem in ultrasound image

Segmentation of the region of interest (Lesion region) is itself a research problem. From an ultrasound image segmenting out the lesion region is very difficult, because there are fatty tissues, shadows, noises of almost similar textures. Defining a robust set of feature vector to ensure maximum classification of false positive instances. Since, classifying a Benign case as Malignant won't harm but on the other hand classifying a Malignant case as a Benign is dangerous. So ROC is needed to be calculated in such way. In case of estimating the strain, large motion is a big problem. Large motion, lateral motion, rotation, non affine motion, these kind of motion is not good for strain estimation by any of the existing methods.

Another important research challenge is the overlapping feature[14] properties of the benign and malignant breast lesion. There can be many different categories that can be found from the ultrasound image. A brief description follows with their overlapping nature.

Features	Explanation
Margin Sharpness	Sharpness of the lesion margin
Margin Echogenicity	Homogenous hypo-echoic pattern with preserved echogenic breast fat tissue mainly observed in benign lumps
Angular Continuity	Continuity of angle in the edge
Saliency	State or quality by which it stands out relative to its neighbors
Entropy	Randomness in the lesion area
FNPA	Computes 4-neighborhood-pixels (second-order) texture parameter of MB within the mask LROI that defines the lesion
Hurst Coefficient	Computes Hurst coefficient that characterizes surface roughness in an image
Margin Definition	Evaluates how well-defined tumor margin is
Aspect Ratio	Computes the aspect ratio of the traced ROI
Solidity	It is the ratio between area and convex area
Convexity and Hausdorf dimension	It is the ratio between convex perimeter and perimeter

Table 1: Features those give best possible distinction between the benign and malignant lesion and Their properties.

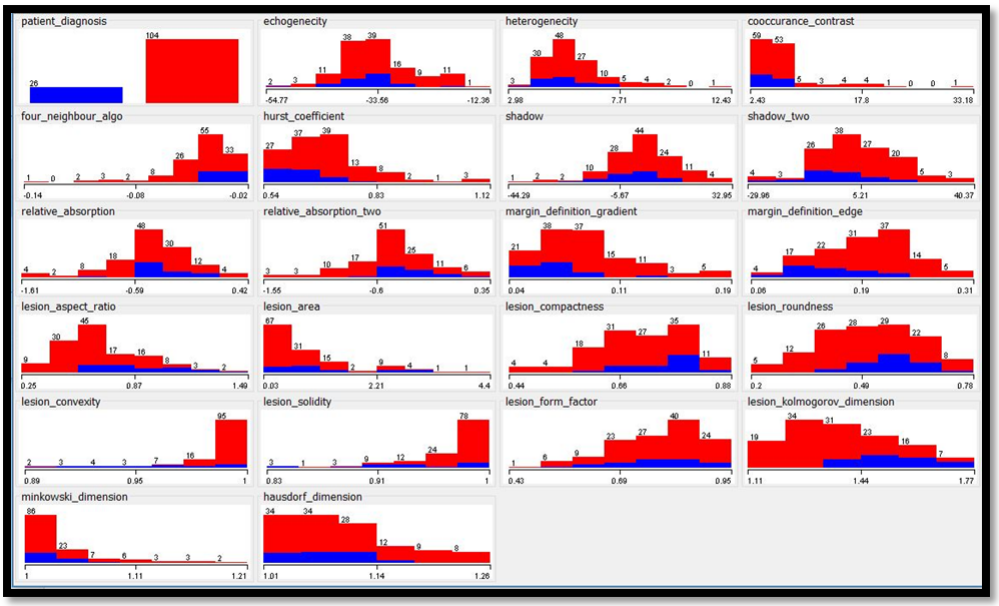


Figure 1.6: Features overlapping. Benign Lesions indicated by red, Malignant ones by blue. Samples are plotted over feature values.

In real patient motion the data acquisition scenario is not ideal and there is movement in lateral and axial direction. So it becomes important to track down the lesion region when strain is applied on overall tissue.

1.4 Thesis Objectives

Our main objective is to figure out the tissue property in order to successfully identify the rigid lesion region through applying pressure on the breast tissue. In general terms, benign lesions compress with transducer pressure and malignant lesions displace the breast tissue without changing in height; this is the basis for our approach.In a nutshell, Tissue rigidity must be clearly depicted from the strain motion.

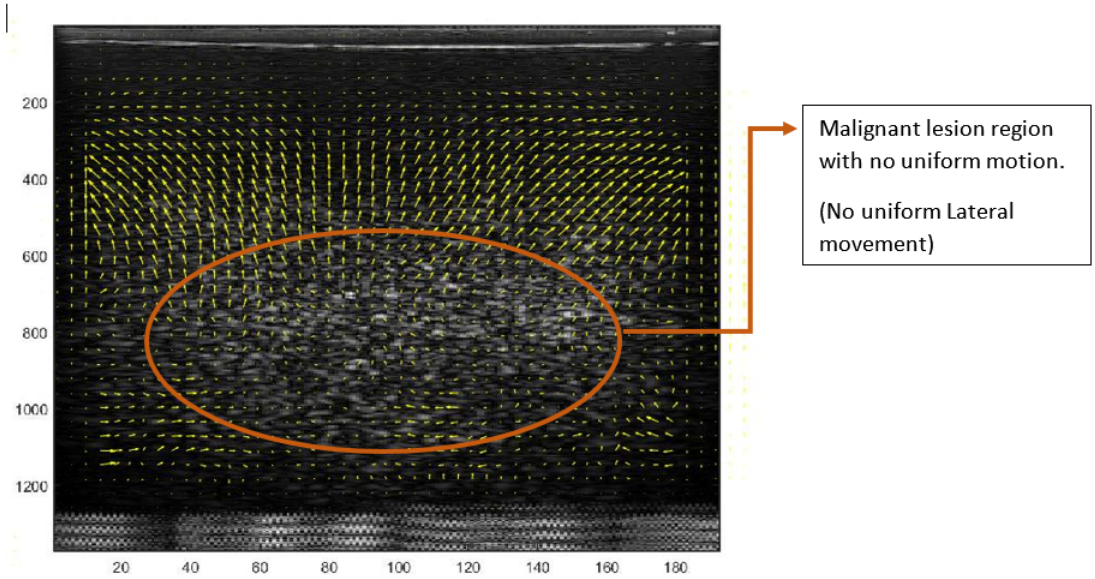


Figure 1.7: Objective : differentiating lesion region effectively from aforementioned tissue property.

1.5 Thesis contributions

- We managed to obtain better performance on higher compression.The result is average when the strain percentage is lower but gives a better result when the strain percentage is increased.
- Physiologically prominent features have been identified. We needed a feature vector which will significantly distinguish between Benign and Malignant Cancerous lesion. We studied that the strain estimation shows significantly different behavior in those two types whereas benign lesions compress with transducer pressure and malignant lesions displace the breast tissue without changing in height. We segment out the uniform stretching region from strain image and that is the malignant lesion.
- Optical flow implementation with median filter over "ultrasound images". Optical flow is a method where we can get the motion vector along x and y both axis by comparing two sequential frames and calculating the gradient from the two frame. By using this method on pre-compression and post-compression frame we can calculate the axial and lateral strain. We examine that before optical flow calculation if median filter applied over the images the performance is improved.

1.6 Organization of the Thesis

The rest of this thesis is organized as follows.

Chapter 2, Gives an overview of different approaches for the strain estimation problem. This chapter also describes the reasons of choosing Lucas Kanade algorithm for Optical flow Method and reasoning Median Filter.

Chapter 3, Proposes a solution to increase accuracy for Strain estimation Problem. This solution contains four separate parts which are fused together to solve the problem. Includes implementation of the proposed methods and also contains an evaluation of proposed methodology.

Chapter 4, Presents Result Analysis and comparison with different studies.

Chapter 5, Presents conclusions and discusses future work.

Chapter 2

2 Literature Review

2.1 Feature Extraction and Performance Analysis

2.1.1 Study of Textural Features

This study[11] presents a computer-aided diagnosis (CAD) system with textural features for classifying benign and malignant breast tumors on medical ultrasound systems. A series of pathologically proven breast tumors were evaluated using the support vector machine (SVM) in the differential diagnosis of breast tumors. The proposed CAD system[11] utilized facile textural features, i.e., block difference of inverse probabilities, block variation of local correlation coefficients and auto-covariance matrix, to identify breast tumor. An SVM classifier using the textual features classified the tumor as benign or malignant. The proposed system[11] identifies breast tumors with a comparatively high accuracy. This can help inexperienced physicians avoid misdiagnosis. The main advantage of the proposed system is that the training and diagnosis procedure of SVM are faster and more stable than that of multilayer perception neural networks. With the expansion of the database, new cases can easily be gathered and used as references. This study[11] dramatically reduces the training and diagnosis time. The SVM is a reliable choice for the proposed CAD system because it is fast and excellent in ultrasound image classification.

They have used the features-

- Block difference of inverse probabilities(BDIP)
- Block variation of local correlation coefficients (BVLC)
- Auto-Covariance Matrix

Accuracy after classification: 93.6%

Also in [12], the study have used following textural features as those gave them the best result in NN Classifier

- Spatial Gray level Dependent Matrices (SGLDM)
- Gray Level Difference Matrices (GLDM)
- Auto Covariance Matrix

2.1.2 Study of Quantitative Features

The goal of this study[10] is to evaluate an Artificial Neural Network (ANN) for differentiating benign and malignant breast masses on ultrasound scans. The ANN was designed with three layers (input, hidden and output layer), where a sigmoidal (hyperbolic tangent) response function is used as an activation function at each unit. Data from 54 patients with biopsy-proven malignant (N=20) and benign (N=34) masses were used to evaluate the diagnostic performance of the ANN. Of the seven quantitative features extracted from ultrasound images, only four showed statistically significant difference between the two categories. These features were margin sharpness, margin echogenicity, angular continuity, and age of patients. The diagnostic performance was evaluated by round-robin substitution to negate bias due to small sample size. All the input features were standardized to zero-mean and unitvariance to prevent non-uniform learning, which can generate unwanted error. The outputs of the network were analyzed by Receiver Operating Characteristics (ROC). The resulting area under the ROC curve Az was 0.856 with 95% confidence limit from 0.734

to 0.936, providing 76.5% specificity at 95% sensitivity. The performance of the ANN was comparable to the performance by logistic regression analysis reported by our group earlier. These results suggest that an ANN when combined with sonography can effectively classify malignant and benign breast lesions. They have used the features- Margin Sharpness, Echogenicity, Angular Continuity, Tissue Attenuation, Mass Attenuation, Excess Attenuation, and Patient Age. They have found that only 4 features show difference. They are- Margin Sharpness, Margin Echogenicity (Relative brightness) Homogenous hypo-echoic pattern with preserved echogenic breast fat tissue mainly observed in benign lumps, Angular Continuity, Patient Age.

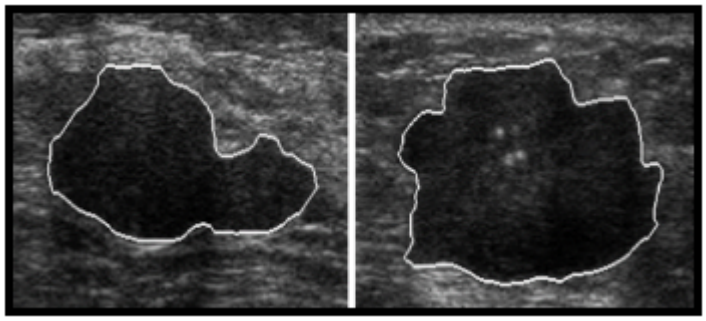


Figure 2.1: Margin Sharpness

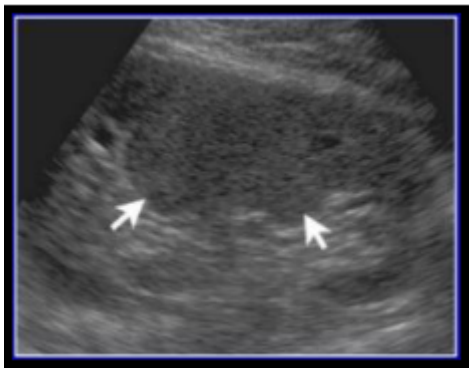


Figure 2.2: Angular Continuty

2.1.3 Study of feature extraction by spatial compounding

Spatial Compounding imaging is a technique in which a number of co-planar, tomographic ultrasound images of an object are obtained from different directions, then combined into a single compound image. With the emergence of recent technology in breast ultrasound, sonographic image quality has changed profoundly. Most notably, the technique of real-time spatial compounding impacts the appearance of lesions and parenchyma. During image acquisition, spatial compounding can be turned on or off at the discretion of the radiologist, but this information is not stored along with the image data. The ability to distinguishing between lesions imaged with and without spatial compounding using either single image features or a Bayesian neural net (BNN) - was assessed using ROC analysis. The database of this study[13] consisted of consecutively collected HDI5000 images of 129 lesions imaged without spatial compounding (357 images, cancer prevalence of 18%) and 370 lesions imaged with spatial compounding (965 images, cancer prevalence 15%). These were used in automated feature selection and BNN training. An additional 33 lesions were imaged for which identical views with and without spatial compounding were available (70 images, cancer prevalence 15%). These

served as an independent test dataset. Lesions were outlined by a radiologist and image features, mathematically describing lesion characteristics, were calculated. In feature selection, the 4 best performing features were related to gradient strength and entropy. The average gradient strength within a lesion obtained an area under the ROC curve (AUC) of 0.78 in the task of distinguishing lesions imaged with and without spatial compounding. The BNN - using 4 features - achieved an AUC on the independent test dataset of 0.98 in this task. The sonographic appearance of breast lesions is affected by spatial compound imaging and lesion features may be used to automatically separate images as obtained with or without this technique. In computer-aided diagnosis (CADx), it will likely be beneficial to separate images as such before using separate classifiers for assessment of malignancy. They have used the features- Gradient Strength Features, Continuous Intensity, Edge Detection (Sobel Mask), Entropy Related Features, Saliency (Relativeness), Entropy, Scale & self-similarity.

2.1.4 Study of Acoustic features

The study[14] have developed quantitative descriptors to provide an objective means of noninvasive identification of cancerous breast lesions. These descriptors include quantitative acoustic features assessed using spectrum analysis of US radio frequency echo signals and morphometric properties related to lesion shape. Acoustic features include measures of echogenicity, heterogeneity and shadowing, computed by generating spectral parameter images of the lesion and surrounding tissues. Spectral-parameter values are derived from rf echo signals at each pixel using a sliding window Fourier analysis. Initial results of biopsy proven cases show that although a single parameter cannot reliably discriminate cancerous from noncancerous breast lesion, multi features analysis provide excellent discrimination for this data set. They had process data for 130 biopsy proven patients acquired during routine US examination at 3 clinical sites and produced an area under the receiver operating characteristic ROC curve of .947. Among the quantitative descriptors, lesion-margin definition, spiculation and border irregularity are the most useful.

Used Features:-
Morphometric features- Aspect Ratio, Margin Definition, Border Irregularity.
Textural features- Echogenicity, Heterogeneity, Shadowing.

Used Classifier:-
Linear-Discriminant Analysis (LDA), Logistic Regression(LR)

Features that have contrasting effect on Benign and Malignant Lesions- FNPA, Hurst Coefficient, Margin Definition, Aspect Ratio, Solidity, Convexity and Hausdorf dimension.

2.1.5 Study of Strain Estimation feature

The basic principles of using sonographic techniques for imaging the elastic properties of tissues are described, with particular emphasis on elastography. After some preliminaries that describe some basic tissue stiffness measurements and some contrast transfer limitations of strain images are presented, four types of elastograms are described, which include axial strain, lateral strain, modulus and Poisson's ratio elastograms. The strain filter formalism and its utility in understanding the noise performance of the elastographic process is then given, as well as its use for various image improvements. After discussing some main classes of elastographic artefacts, the paper [14]concludes with recent results of tissue elastography in vitro and in vivo.

2.2 Classifier Performance Study

2.2.1 Neural Network

1. Artificial Neural Network: The goal of this study [10] is to evaluate an Artificial Neural Network (ANN) for differentiating benign and malignant breast masses on ultrasound scans.

The ANN was designed with three layers (input, hidden and output layer), where a sigmoid (hyperbolic tangent) response function is used as an activation function at each unit. Data from 54 patients with biopsy-proven malignant (N=20) and benign (N=34) masses were used to evaluate the diagnostic performance of the ANN. Of the seven quantitative features extracted from ultrasound images, only four showed statistically significant difference between the two categories. These features were margin sharpness, margin echogenicity, angular continuity, and age of patients. The diagnostic performance was evaluated by round-robin substitution to negate bias due to small sample size. All the input features were standardized to zero-mean and unit-variance to prevent non-uniform learning, which can generate unwanted error. The outputs of the network were analyzed by Receiver Operating Characteristics (ROC). The resulting area under the ROC curve Az was 0.856 with 95% confidence limit from 0.734 to 0.936, providing 76.5% specificity at 95% sensitivity. The performance of the ANN was comparable to the performance by logistic regression analysis reported by our group earlier. These results suggest that an ANN when combined with sonography can effectively classify malignant and benign breast lesions.

2. Generalized Neural Network: They [12] have used textural features as it gives them the best result in NN Classifier.
   Used Textural Features- Spatial Gray level Dependent Matrices (SGLDM), Gray Level Difference Matrices(GLDM), Auto Covariance Matrix.

fn*	m,n	Accuracy(%)
3	3x3	82.38
	5x5	81.67
	7x7	83.62
5	3x3	82.29
	5x5	84.38
	7x7	82.95
7	3x3	84.23
	5x5	82.33
	7x7	84.33

Table 2: Varying the size of m*n matrices the accuracy varies. (*fn = Number of ROI sub-images to extract features)

3. Bayesian Neural Network: With the emergence of recent technology in breast ultrasound, sonographic image quality has changed profoundly. Most notably, the technique of real-time spatial compounding [13] impacts the appearance of lesions and parenchyma. During image acquisition, spatial compounding can be turned on or off at the discretion of the radiologist, but this information is not stored along with the image data. The ability to distinguishing between lesions imaged with and without spatial compounding - using either single image features or a Bayesian neural net (BNN) - was assessed using ROC analysis. Our database consisted of consecutively collected HDI5000 images of 129 lesions imaged without spatial compounding (357 images, cancer prevalence of 18%) and 370 lesions imaged with spatial compounding (965 images, cancer prevalence 15%). These were used in automated feature selection and BNN training. An additional 33 lesions were imaged for which identical views with and without spatial compounding were available (70 images, cancer prevalence 15%). These served as an independent test dataset. Lesions were outlined by a radiologist and image features, mathematically describing lesion characteristics, were calculated. In feature selection, the 4 best performing features were related to gradient strength and entropy. The average gradient strength within a lesion obtained an area under the ROC curve (AUC) of

0.78 in the task of distinguishing lesions imaged with and without spatial compounding. The BNN - using 4 features - achieved an AUC on the independent test dataset of 0.98 in this task. The sonographic appearance of breast lesions is affected by spatial compound imaging and lesion features may be used to automatically separate images as obtained with or without this technique. In computer-aided diagnosis (CADx), it will likely be beneficial to separate images as such before using separate classifiers for assessment of malignancy.They have used the features- Gradient Strength Features, Continuous Intensity, Edge Detection (Sobel Mask) Entropy Related Features- Saliency (Relativeness), Entropy,Scale & self-similarity.

2.2.2 Support Vector Machine

The study of[14] presents a computer-aided diagnosis (CAD) system with textural features for classifying benign and malignant breast tumors on medical ultrasound systems. A series of pathologically proven breast tumors were evaluated using the support vector machine (SVM) in the differential diagnosis of breast tumors. The proposed CAD system utilized facile textural features, i.e., block difference of inverse probabilities, block variation of local correlation coefficients and auto-covariance matrix, to identify breast tumor. An SVM classifier using the textual features classified the tumor as benign or malignant. The proposed system identifies breast tumors with a comparatively high accuracy. This can help inexperienced physicians avoid misdiagnosis. The main advantage of the proposed system is that the training and diagnosis procedure of SVM are faster and more stable than that of multilayer perception neural networks. With the expansion of the database, new cases can easily be gathered and used as references. This study dramatically reduces the training and diagnosis time. The SVM is a reliable choice for the proposed CAD system because it is fast and excellent in ultrasound image classification. They have used the features- Block difference of inverse probabilities(BDIP), Block variation of local correlation coefficients (BVLC), Auto-Covariance Matrix.

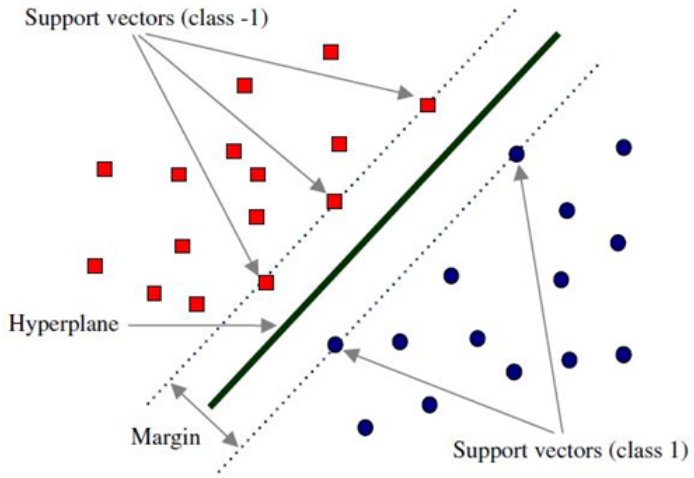


Figure 2.3: Hyper plane in SVM and margin

2.2.3 Linear Discriminant Analysis & Logistic Regression

The study of [14] developed quantitative descriptors to provide an objective means of noninvasive identification of cancerous breast lesions. These descriptors include quantitative acoustic features assessed using spectrum analysis of US radio frequency echo signals and morphometric properties related to lesion shape. Acoustic features include measures of echogenicity, heterogeneity and shadowing, computed by generating spectral parameter images of the lesion and surrounding tissues. Spectral-parameter values are derived from rf echo signals at each pixel using a sliding window

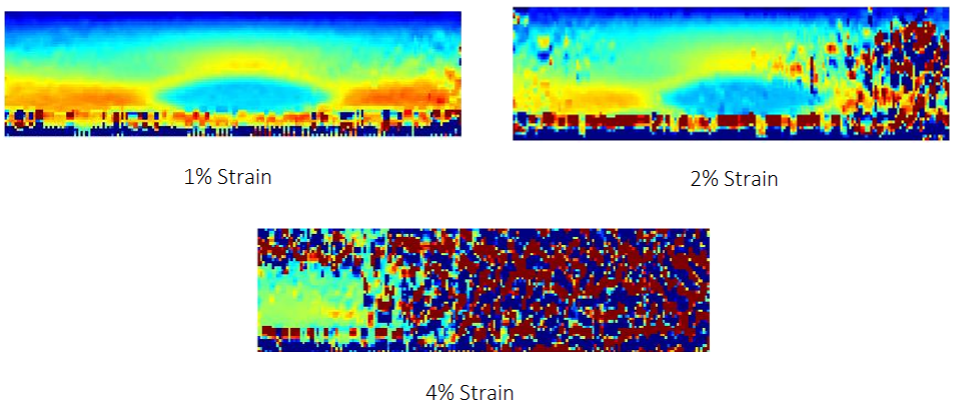


Figure 2.4: Performance of this method decreases as the strain increase.[5]

Fourier analysis. Initial results of biopsy proven cases show that although a single parameter cannot reliably discriminate cancerous from non-cancerous breast lesion, multi features analysis provide excellent discrimination for this data set. They had process data for 130 biopsy proven patients acquired during routine US examination at 3 clinical sites and produced an area under the receiver operating characteristic ROC curve of .947. Among the quantitative descriptors, lesion-margin definition, spiculation and border irregularity are the most useful. Features that have contrasting effect on Benign and Malignant Lesions- FNPA, Hurst Coefficient, Margin Definition, Aspect Ratio, Solidity, Convex, ity and Hausdorf dimension.

2.3 Strain Estimation Study

In strain estimation study [5] it has been said that cross correlation between pre and post compression RF data to estimate local displacement is calculated for estimating the strain. Result is improved by using smooth windows (To reduce noise). SNR and CNR significantly improved in this method for estimating the strain information. But in this method, the performance is decreased as the strain is increased.

2.4 Optical Flow based approaches on Radio Frequency Signals

As the effectiveness of an early detection of breast cancer using the mammography method alone is uncertain, it is crucial to provide an alternative method instead. This paper[3] analyzes two optical flow algorithms utilizing a gradient method to aid current imaging techniques for a potential alternative method in aiding early breast cancer detection. The gradient method is a cost effective method that has the potential to be a mass screening method for this purpose. This paper compares two optical flow algorithms that are capable to detect the motion of breast tumor on B-mode ultrasound images. An analysis of 2D images of breast cancer lesions are compared using two radient optical flow algorithms: Horn & Schunck and Lucas & Kanade. Both algorithms successfully show the direction of the tumor motion. However, while Lucas & Kanade can handle the short motion displacement of the tumor on the tested ultrasound images, Horn & Shunck failed to do so. This implies that the Lucas & Kanade algorithm is potentially more effective in handling ultrasound images of breast tumor. The results obtained showed that the Lucas & Kanade give better accuracy compared to Horn & Schunk.
Optical flow is used in detecting motion of lesion and displacement. There are two methods for optical flow on B-mode Image-

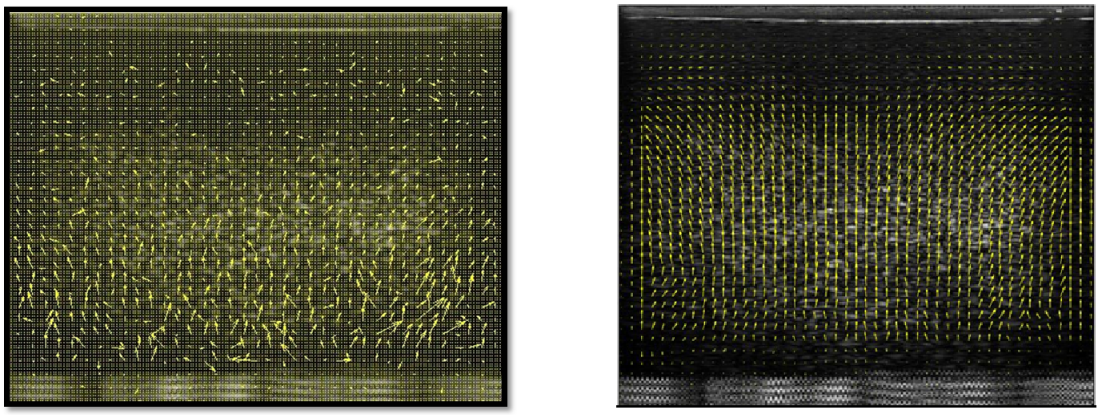


Figure 2.5: Motion vector using (a) Horn Shunck and (b) Lucas Kanade

1. Horn Shunck
2. Lucas Kanade

Lucas Kanade gives better result for small displacement

Optical flow (OF) method has been used in ultrasound elastography to estimate the strain distribution in tissues. However the bias of strain estimation by OF has previously been shown to be close to 20%. The objective in this paper[4] is to improve the performance of OF-based strain estimation, a two-step OF method with a local warping technique is proposed in this paper. The local warping technique effectively decreases the decorrelation of the signals, and hence improves the performance of strain estimation. Simulations on both homogeneous and heterogeneous models with different strains are performed. Experiments on a heterogeneous tissue-mimicking phantom are also carried out. Simulation results of the homogeneous model show that the two-step OF method reduces the bias of strain estimation from 23.77% to 1.65%, and reduces the standard deviation of strain estimation from 2.9 10-3 to 0.47 10-3. Simulation results of the heterogeneous model shows that the signals-to-noise ratio (SNRe) of strain estimation is improved by 2.1 and 5.3 dB in the inclusion and background, respectively, and the contrast-to-noise ratio (CNRe) is improved by 6.8 dB. Finally, results of phantom experiments show that, by using the proposed method, the SNRe is increased by 4.0 dB and 8.9 dB in the inclusion and background, respectively, while the CNRe is increased by 13.1 dB. The proposed two-step OF method is thus demonstrated capable of improving the performance of strain estimation in OF-based elastography.

- The first step of the two-step OF method is the conventional OFmethod.
- In the second step, the local warping technique is performed. The axial strain and axial shear strain estimated in the first step are used to warp the ROI from the pre-deformed RF signals.
- The OF method is implemented again to estimate the full strain tensors.
- The final axial strains are calculated from the combination of the axial strains estimated in steps 1 and 2.

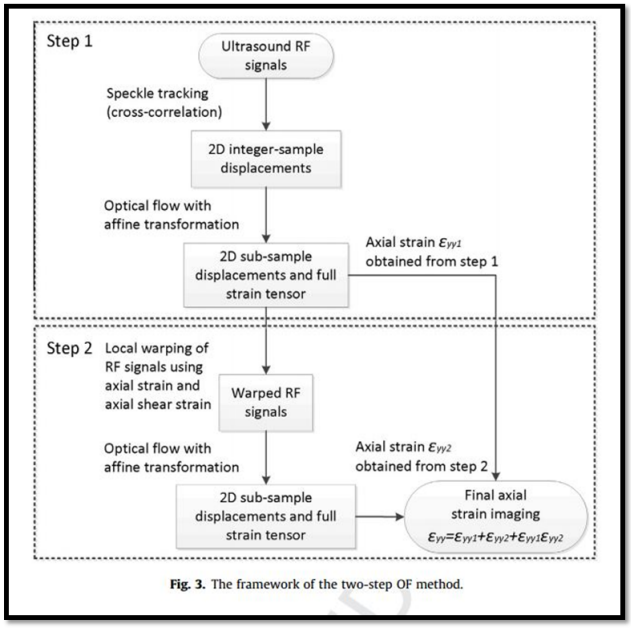


Figure 2.6: Two step optical Flow calculation

$$\varepsilon_{yy} = \varepsilon_{yy_1} + \varepsilon_{yy_2} + \varepsilon_{yy_1} * \varepsilon_{yy_2}$$

Here, $\varepsilon_{yy_1} =$ axial strain estimated from initial image, $\varepsilon_{yy_1} =$ axial strain estimated from the warped pre-deformed, and ε_{yy} is the finally calculated optical flow

Chapter 3

3 Proposed Method

3.1 Our Proposal And Building Block

Our proposal and methodologies concern with the accuracy and visualization of the lesion region that is affected by the pressure or strain applied on overall breast tissue. The following points have been depicted from the studies of chapter 2.

1. Since image strain estimation is more accurate for Lucas Kanade Algorithm for optical flow we choose this algorithm for estimating the strain.[3]
2. Optical flow while estimating strain ,finds motion vector for all points and it can easily be interpolated, where Cross Correlation method needs to exclude many intermediate data points when finding strain. Motion vector for maximum data point needs to be found.
3. Median Filter and Smoothing filter applied for the clarity of edge and noise removal in Alam et al [14]. To get competitive result with two step optical method (Xiaochang et al[4]) Median Filter filter can result in good SNR measures.

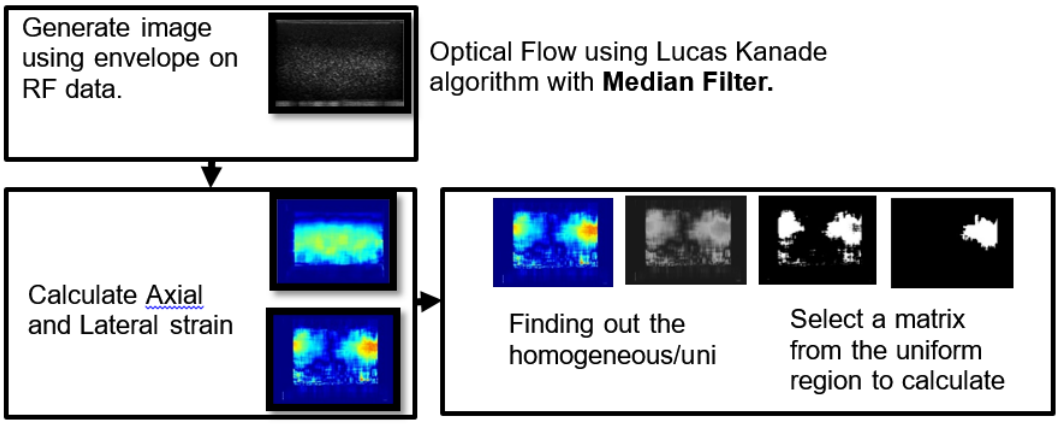


Figure 3.1: Building Block of Our Proposal

Step-1: Generate image using envelope on RF data.

We will be working with scaled image data since the information in the RF data signal is redundant and the motion vector to be identified has to be pictorial to follow the BIRADS convention[12].

Step-2: Optical Flow using Lucas Kanade algorithm.[3]

Step-3: Calculate axial and lateral strain matrix from optical flow.

Step-4: Finding out the homogeneous/uniform stretching region in axial and lateral direction of Y axis.

Step-5: Select a matrix from the uniform region to calculate SNR.

3.2 Generate image using envelop on RF data

Envelop of the RF echo signal is calculated by using the Hilbert transform algorithm. In the envelop generation, Envelope(rf,n) uses n-point FFT to reduce computation time. An advantage of the Hilbert algorithm is that being an asynchronous detector, it does not need the center frequency.

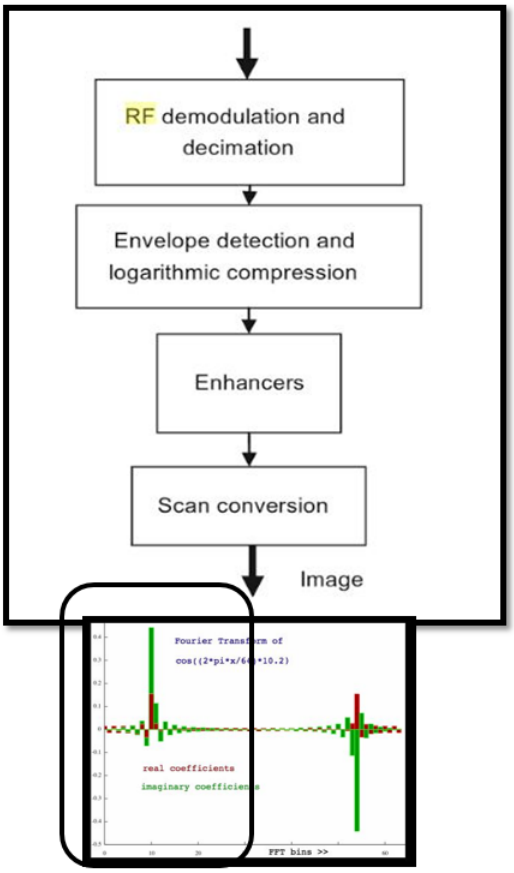
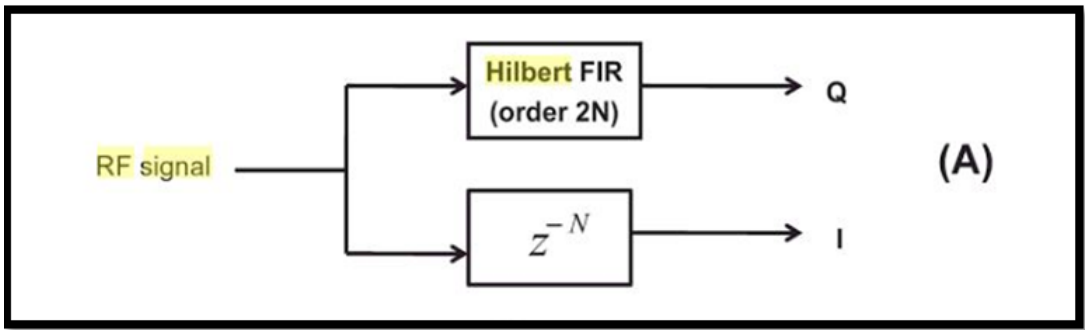


Figure 3.2: RF data to image generation procedure



By using hilbert's algorithm envelop is generated and image is formed from it.

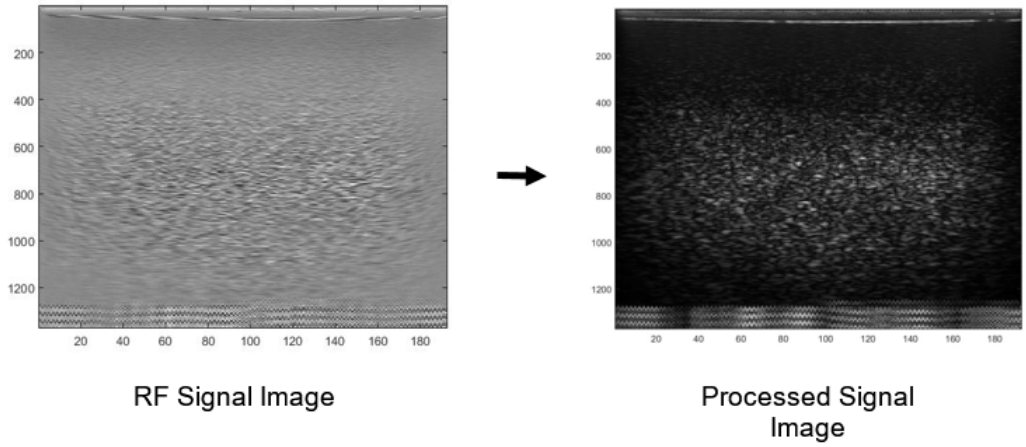


Figure 3.3: RF data and corresponding image

3.3 Optical flow using Lukas Kanade Algorithm

How to estimate pixel motion from image H to image I is shown bellow-

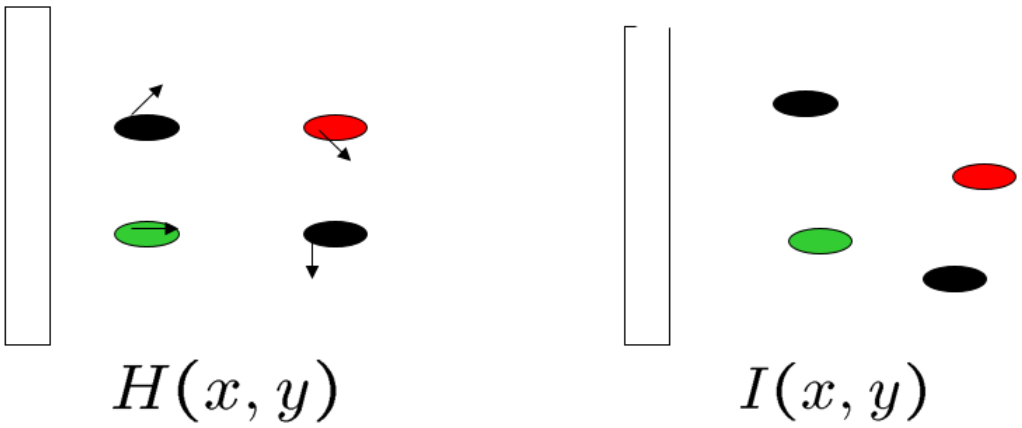


Figure 3.4: Lucas Kanade solves pixel correspondence problem given a pixel in H, look for nearby pixels of the same intensity in I

Key assumptions

Color Constancy: a point in H looks the same in I. For gray scale images, this is **brightness constancy**.
Small Motion: points do not move very far.

This is the **optical flow** problem.

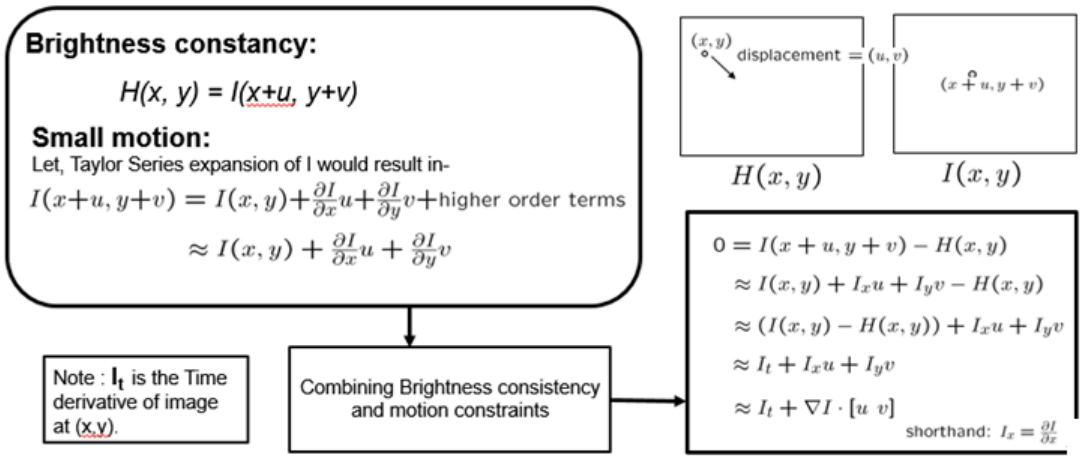


Figure 3.5: Optical Flow problem

Equations and Unknowns per pixel 1 equation, but 2 unknowns (u and v) Intuitively constraint meaning The component of the flow in the gradient direction is determined The component of the flow parallel to an edge is unknown.
Basic idea is Assume motion field is smooth. Lukas & Kanade: Assume locally constant motion Pretend the pixel's neighbors have the same (u,v). If we use a 5x5 window, that gives us 25 equations per pixel.

$$O = I_t + \nabla I.[u.v]$$

Assume motion field is smooth,

Lukas & Kanade:

Assume locally constant motion. Pretend the pixel's neighbors have the same (u, v)
If we use a 5x5 window, that gives us 25 equations per pixel.

$$O = I_t(p_i) + \nabla I(p_i).[uv]$$

$$\begin{bmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_2) & I_y(p_2) \\ \vdots & \vdots \\ I_x(p_25) & I_y(p_25) \end{bmatrix} \times \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(p_1) \\ I_t(p_2) \\ \vdots I_t(p_25) \end{bmatrix}$$

$$A(25 \times 2) \times d(2 \times 1) = b(25 \times 1)$$

Minimum least squares solution given by solution (in d) of

$$(A^T A)d = A^T b$$

3.4 Axial and lateral strain calculation

$$\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} \times \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}$$

The summations are over all pixels in the $k \times k$ window This is solvable when - Firstly $A^T A$ should be invertible and $A^T A$ should not be too small due to noise, eigenvalues λ_1 and λ_2 of $A^T A$ should not be too small

3.5 Finding out homogenous stretching region:

For finding out the homogeneous or uniform region we followed some steps. Firstly, we have shown the strain image using colormap (blue for 0 and red for 255). Then convert the strain image to grayscale image. Next the homogeneous/uniform region is brighter. Then we select a threshold that will differentiate the brighter region. The next step is from the binary image we find out the largest connected component. This largest connected uniform region is selected.

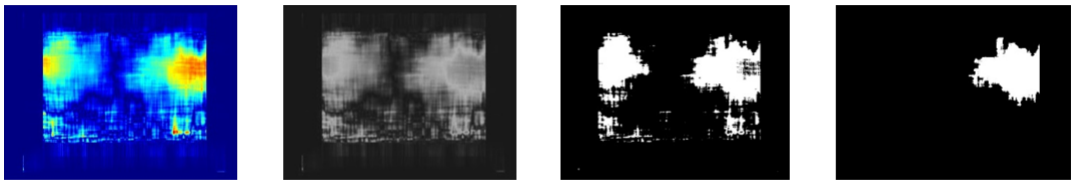


Figure 3.6: Steps of findind homogeneous region

Chapter 4

4 Experimental Analysis

4.1 Dataset

4.1.1 Description

Generally, when collecting data from a live patient,important conditions to keep in mind are the focal point of the transducer and the position from which the data is collected.But standardizing constant pressure is not possible when extracting real patient data.So we normalize our data-set to phantom unit tissue where a solid lesion is artificially created.And then pressure is applied from one of the four sides and other sides remain static.

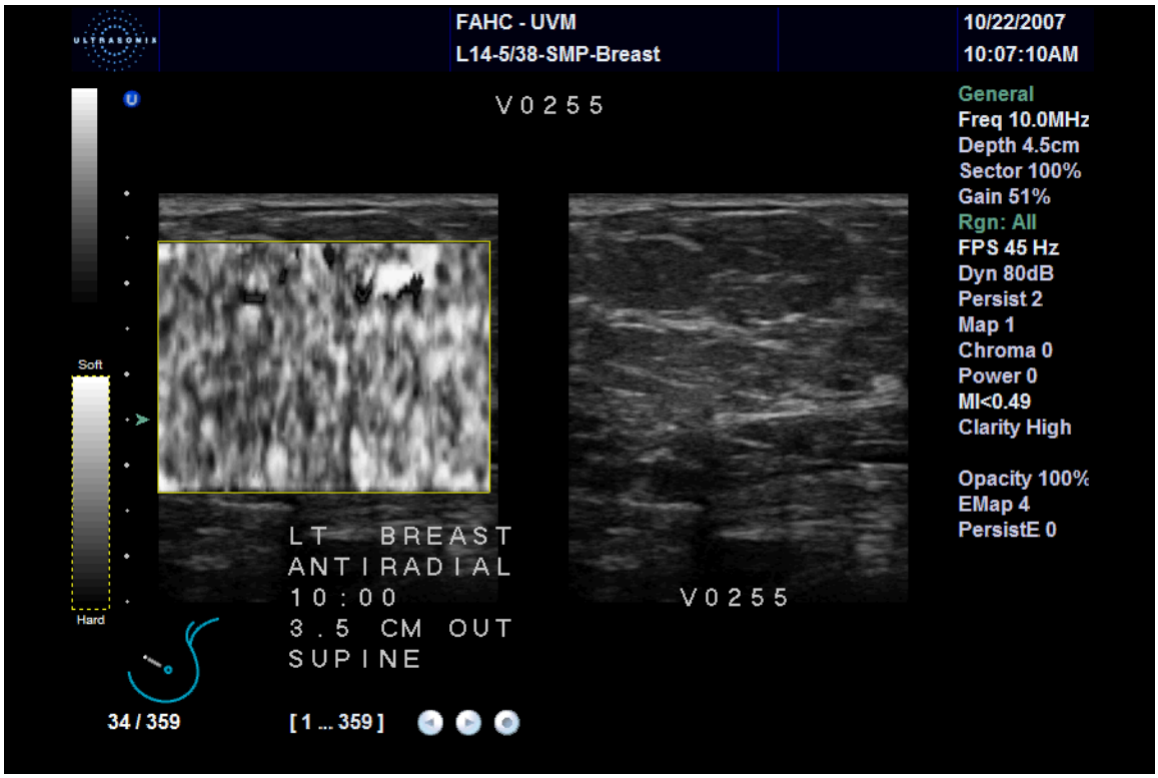
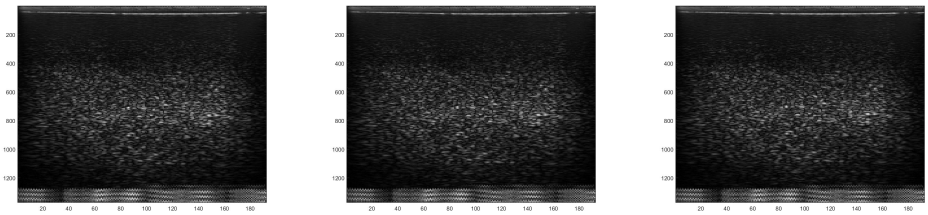


Figure 4.1: Real patient data acquisition.

4.1.2 Experimental Setup

The system that has been used to acquire the data is ATL Ultrasound Scanner with L10-5 probe.Presssure has been applied with mechanical jig mounted on motor unit. Each compression step compressed the phantom unit of tissue by 1%. Data has been collected using ATL in trigger-mode. The focus of transducer was at 2.8 cm.

Step by step percentage-wise compression with the help of mechanical jig.



Filename	Compression steps
(a) AAAAA1-8.eye	Position 1, 0% compression
(b) AAAAB1-8.eye	Position 1, 0% compression
(c) AAAAC1-8.eye	Position 2, 1% compression
(d) AAAAD1-8.eye	Position 3, 2% compression
(e) AAAAE1-8.eye	Position 4, 3% compression
(f) AAAAF1-8.eye	Position 5, 4% compression
(g) AAAAG1-8.eye	Position 6, 5% compression

Figure 4.2: The dataset and their compression percentages with respect to the initial position.

4.2 Performance based on feature evaluation

4.2.1 Feature Performances

We use different classifiers over a data set containing patient data of benign and malignant lesion using a feature vector based on textural features. On individual feature we also use the classifiers to see that how relevant the feature is to distinguish between the two types of cancerous lesion. Along with the classifiers the feature's performance can be measured in group and also distinctly. The distinct performance have been elaborated in the figure given below

4.2.2 Performance Observation

For the performance measurement of different classifiers we use different classifiers over the feature vector and saw the results. The result comparison of different classifiers are shown below.

Classifiers	Correctly Classified instances	Incorrectly classified instances	ROC
Logistic Regression	83.8462	16.1538	0.889
Multi Layer Perceptron (Neural Network	88.4615	11.5385	0.874
Sequential Minimal Optimization (SMO)	89.2308	10.7692	0.774
Lazy LWL	89.2308	10.7692	0.857
K -Star	82.3077	17.6923	0.822
Decision Table	89.2308	10.7692	0.83
Support Vector Machine (SVM)	80	20	0.558
Ada-Boost Classifier	91.5385	8.4615	0.915

Table 3: Different Classifier's Performance

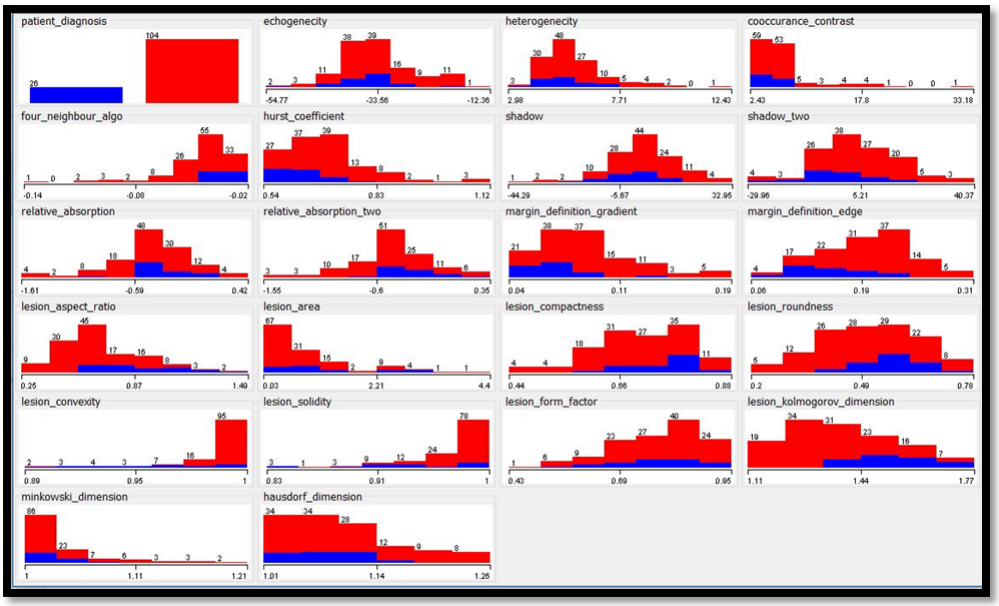


Figure 4.3: Features individual performance. The Red Region depicts the Benign cancer lesion and the Blue one as Malignant

We found that-

- In order to improve classifier result, correlated features must be removed which over-fits the data. Over-fitting occurs when there are more biased features than usual which are more or less similar.These features are removed to ensure that the classification would not has to deal with extra similar information while classifying.
- The figure 2.1 shows the Benign(red) and Malignant(Blue) occurrences over different values of all 21 features.
- This represents which feature is having more classification impact.
- Classifying with individual features we have identified the best features fitting the classification.
- Though some textural features from ultrasound images can give good results using non linear classifier but it is not enough for classifying perfectly. So we need a feature that can give significant difference in those two types. Tissues elastic property is such a feature. By estimating strain by applying pressure we can distinguish between soft tissue and lesion.

4.3 Performance Analysis based on Strain Estimation

4.3.1 Performance Metric

SNR: There are several options (and definitions) for the calculation of the SNR of or in an image. First, there are two incompatible definitions of the SNR: SNR is frequently (e. g., in many engineering applications) defined as the ratio of the signal power and the noise power (which is consistent with the definition by Gonzalez given above), but – particularly in imaging – an alternative definition can be found, where SNR is given as the ratio of the mean value of the signal and the standard deviation of the noise. The power ratio (first definition above) is frequently expressed in dB (using the logarithm), while the signal ratio (second definition) is more often given as a number (of dimension 1).

Second, you must define in which part of the image the signal (power or mean value) is determined. Typical choices are: (1) the maximum power or intensity within the image; this gives you the peak-signal-to-noise ratio (PSNR); (2) the mean power or intensity; or (3) the power or signal of a reference structure within the image (e. g., in medical images with large amounts of (zero signal) background this is more useful than including the background into the mean power or signal).

If the noise standard deviation (and the statistical distribution of noise and its spatial distribution noise may be distributed non-uniformly over the image) is already known, then it's done and SNR can be calculated. (If the noise level is unknown as of in our case (standard deviation or variance), it can be difficult to measure it reliably in an image – at least, if you do not have a noise-free image.)

$$\mathrm{SNR} = \frac{\mu_{\mathrm{sig}}}{\sigma_{\mathrm{bg}}}$$

This is the formula for calculating SNR value for single image

Figure 4.4: Uniform region found

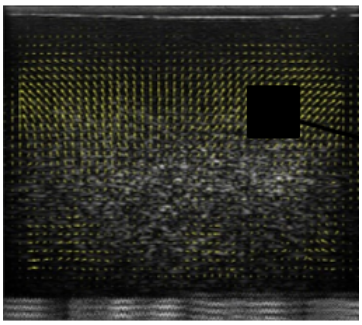


Figure 4.5: SNR is calculated from the uniform region

After finding out the uniform/homogeneous stretching region, we select a kxk matrix for calculating SNR.

4.3.2 Performance on Ultrasound Image

Performance is measured based on Signal to Noise Ratio (SNR) of uniform region and background. lesion movement is minimal among all the soft tissues. when pressure is applied all other soft tissues deform but the lesion remain solid and move a little bit in the same direction of pressure without deformation. There can be two cases possible for lesion movement when pressure is applied. Pressure is always applied in y direction.

1. Case 1: Malignant lesion region with no uniform motion
   In this case, when pressure is applied from downward or in y direction the malignant lesion region has no uniform motion. So we calculate the SNR value from the mean of the uniform region and standard deviation from the background. We compare our result with the method of cross correlation on RF data. The result comparison graph of the cross correlation method and optical flow method is shown bellow. The Table 4 contains the SNR values of two methods in different strain level

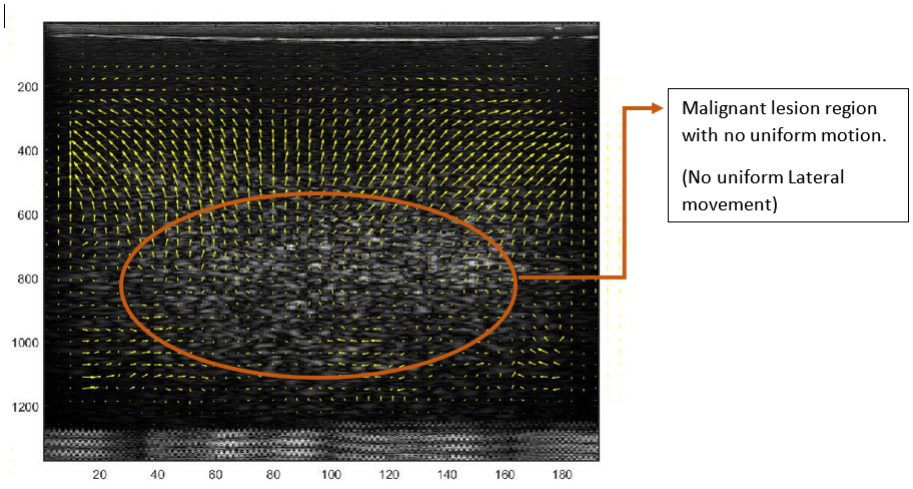


Figure 4.6: Malignant lesion region with no uniform motion

Observation: SNR value is better in optical flow strain estimation than the SNR value which is estimated from cross correlation between RF data. As the strain increased the performance of previous method decreases but in optical flow method it is still better.

	Strain	1%	2%	3%	4%
Cross Correlation	SNR(dB)	20.92	21.04	21.79	18.02
	Mean	0.0111	0.0218	0.0318	0.0431
Optical Folw	SNR(dB)	**23.10**	**26.05**	**27.61**	**26.14**
	Mean	0.3525	0.5445	0.5151	0.4992

Table 4: SNRdB and Mean values of Cross Correlation and Optical Flow method for Case:1

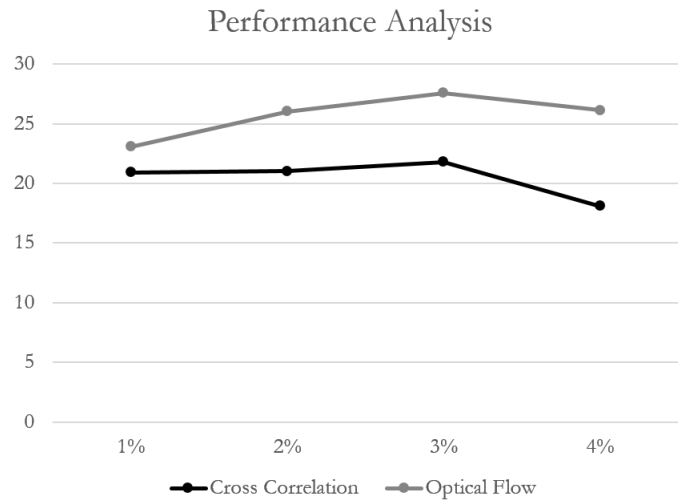


Figure 4.7: SNR_{dB} VS Strain of Cross Correlation and Optical Flow method for Case 1

2. Case2: Uniform upward motion for the malignant lesion
In this case, when pressure is applied from downward or in y direction the malignant lesion region has upward uniform motion. We calculate the SNR value from the mean of the uniform region(not the lesion uniform region) and standard deviation from the background. We compare our result with the method of cross correlation on RF data. The result comparison graph of the cross correlation method and optical flow method is shown bellow. The Table 5 contains the SNR values of two methods in different strain level.

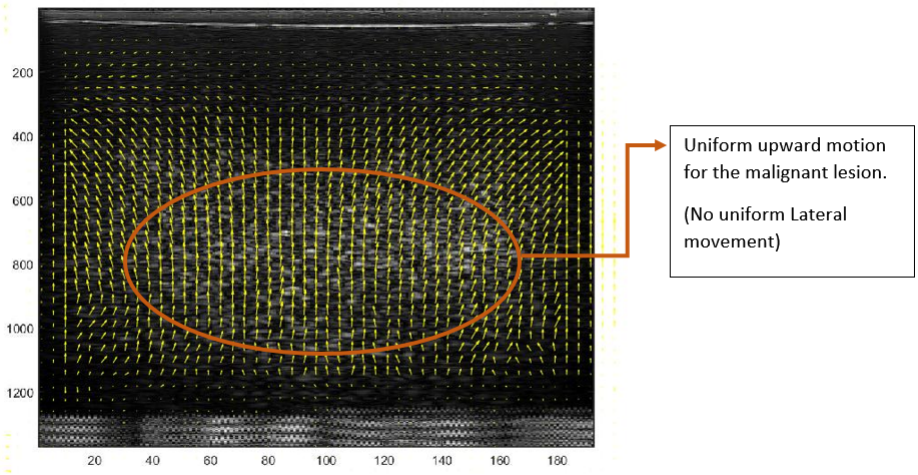


Figure 4.8: Uniform upward motion for the malignant lesion

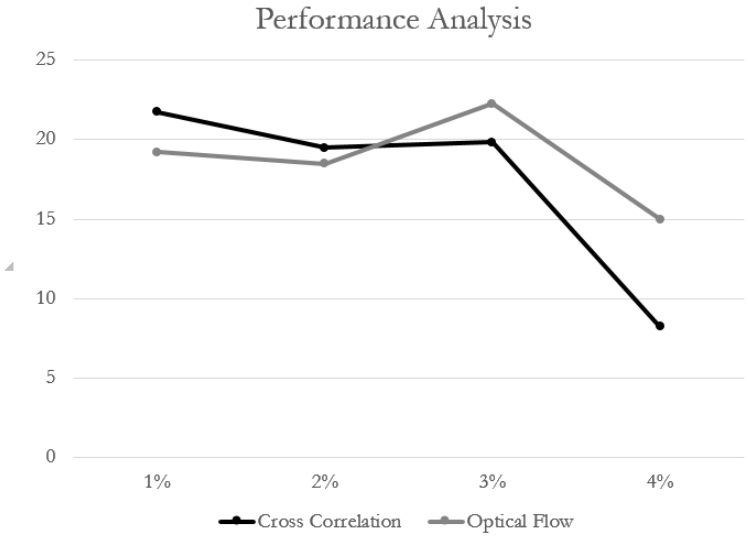


Figure 4.9: SNR_{dB} VS Strain of Cross Correlation and Optical Flow method for Case 2

	Strain	1%	2%	3%	4%
Cross Correlation	SNR(dB)	**21.74**	**19.47**	19.82	8.24
	Mean	0.0101	0.0201	0.0299	0.0297
Optical Folw	SNR(dB)	19.02	18.48	**22.26**	**14.97**
	Mean	0.5598	0.5144	0.5057	0.4848

Table 5: SNR_{dB} and Mean values of Cross Correlation and Optical Flow method for Case:2

Observation: SNR value is better in optical flow strain estimation when the strain is increased than the SNR value which is estimated from cross correlation between RF data.

Chapter 5

5 Conclusion

Strain estimation by optical flow method on images give better results than cross-correlation technique on RF data. Optical flow method results in better accuracy even with more than 2% strain. The results are competitive with Two step Optical flow technique that has also been done over RF data. From the image itself motion estimation is useful since the motion can be evaluated using the handful of techniques which is much better than using the RF data directly.As for our proposed method it can be used as a segmentation technique for the lesion region using optical flow.After segmentation we can use warp technique to improve and experiment further.Using different optimization techniques of optical flow,iterative optical flow estimation,robust motion estimation,global smoothing,probabilistic formulation for further improvement can result in a better outcome for the strain estimation technique.

References

[1] J Ophir, S K Alam, B Garra, F Kallel, E Konofagou, T Krouskop and T Varghese, Elastography: ultrasonic estimation and imaging of the elastic properties of tissues

[2] R. Ahmed, R. Afrin, M. Rubel, K. Islam, C, Jia, D, Metaxas, B. Garra, S. Alam, Comparison of windowing effects on elastography images: Simulation phantom and in vivo studies, Ultrasonics 66(2016)140-153

[3] F. Shuib, M. Othman, K. Abdulrahim, and Z. Zulkifli, Analysis of Motion Detection of Breast Tumor Based on Tissue Elasticity From B-Mode Ultrasound Images Using Gradient Method Optical Flow Algorithm, 978-1-4799-3251-1/13

[4] X. Pan, J. Gao, S. Tao, J. Bai, J. Luo, A two-step optical flow method for strain estimation in elastography Simulation and phantom study, http://dx.doi.org/10.1016/j.ultras.2013.11.010

[5] S. Kaisar Alam, Er Nest J. Feleppa, Mark Ron Deau, Andrew Kalisz Brian S. Garra Ultrasonic Multi-Feature Analysis Procedure for Computer-Aided Diagnosis of Solid Breast Lesions. ultrasonic imaging 33,17-38 (2011)

[6] L. Gao, K. J. Parker, R. M. Lerner and S. F. Levinson, Imaging of the elastic properties of tissue-a review, Ultrasound in Med. & Biol., Vol. 22. No. 8. pp. 959-977. 1996

[7] A. Kumar, J. Ophir and Thomas A. Krouskop, Noise Performance and Signal-To-Noise Ratio of Shear Strain Elastograms, ultrasonic imaging 27, 145-165 (2005)

[8] Jonathan Porée, Damien Garcia, Boris Chayer, Jacques Ohayon, and Guy Cloutier, Noninvasive Vascular Elastography With Plane Strain Incompressibility Assumption Using Ultrafast Coherent Compound Plane Wave Imaging, ieee transactions on medical imaging, vol. 34, no. 12, december 2015

[9] Jonathan Ophir, Faouzi Kallel, Tomy Varghese, Michel Bertrand, Ignacio Ce' spedes, Hari Ponnekanti, Elastography: A Systems Approach

[10] 1W.Shen,R.Chang,W.Moon,Y.Chou,C.Huang,Breast Ultrasound Computer Aided Diagnosis Using BI-RADS Features, Academic Radiology 14(2007)928-939

[11] Y.Hunang, K.Wang, B.Chen, diagnosis of breast tumors with ultrasonic texture analysis using support vector machines. Neural computing and applications 15 (2)(2006) 164-169

[12] 12. W.Shen,R.Chang,W.Moon,Y.Chou,C.Huang,Breast Ultrasound Computer Aided Diagnosis Using BI-RADS Features ,Academic Radiology 14(2007)928-939

[13] 13. K.Drukker, C.A.Sennett, M.L.Giger, The effect of image quality on the appearance of lesions on breast ultrasound : implications for CADx, in : Proceedings of SPIE, Medical Imaging 2007 : Computer – Aided Diagnosis, vol. 6514, 2007,p.65141E

[14] 14. S. Kaisar Alam, Er Nest J. Feleppa, Mark Ron Deau, Andrew Kalisz Brian S. Garra Ultrasonic Multi-Feature Analysis Procedure for Computer-Aided Diagnosis of Solid Breast Lesions. ultrasonic imaging 33,17-38 (2011)

[15] 15. A novel automatic seed point selection algorithm for breast ultrasound images

[16] 16. R.Jemila Rose,S.Allwin Computerized Cancer Detection and Classification Using Ultrasound Images: A Survey International Journal of Engineering Research and Development e-ISSN: 2278-067X, p-ISSN : 2278-800X, www.ijerd.com Volume 5, Issue 7 (January 2013), PP. 36-47

[17] 17. Tao Tan, Jan-Jurre Mordang, Jan van Zelst Computer-aided detection of breast cancers using Haar-like features in automated 3D breast ultrasound. Medical Physics 42, 1498 (2015); doi: 10.1118/1.4914162

[18] 18. Karen Drukker, PhD Maryellen L. Giger, PhD Charles E. Metz, PhD Robustness of Computerized Lesion Detection and Classification Scheme across Different Breast US Platforms. 10.1148/radiol.2373041418Radiology 2005

[19] 19. Juan Shan,s. Kaisar Alam, Brian Garra, Yingtao Zhang, Tahira Ahmed Computer-aided Diagnosis For Breast Ultrasound Using computerized Bi-rads Features And Machine Learning Methods. Ultrasound in Med. & Biol., Vol. 42, No. 4, pp. 980–988, 2016

[20] 20. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics 2014. CA Cancer J Clin.2014;64:9–29