1998

Wavelet-Based Registration of Medical Images.

Raj Sharman
Louisiana State University and Agricultural & Mechanical College

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

WAVELET BASED REGISTRATION OF MEDICAL IMAGES

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Computer Science

by
Raj Sharman
B.Tech., Indian Institute of Technology, 1980
M.Tech., Indian Institute of Technology, 1983
M.S., Louisiana State University, 1987
August, 1998

ACKNOWLEDGEMENTS

My advisor Dr. John M. Tyler has made great sacrifices for me. I am grateful to him for all he has done for me both on a personal and professional level. He has been more than an advisor. Having worked with Dr. John Tyler, I have come to appreciate him in many ways and his way of looking at things. Dr. Tyler's guidance in my research has helped me immensely. I truly appreciate the time he took to read my dissertation. I will always remember that he spent long hours the entire Christmas holidays of 1997 working with me. I would also like to thank his wife, Cindy for being very accommodating.

I also wish to express my deep gratitude and appreciation to Dr. Kundu for his suggestions that has helped improve this dissertation. He has also been a source of great encouragement to me. On several occasions, he has asked me probing questions that has lead to the improvement of this research. He has helped me grow and better myself. I am grateful to him for this experience and also for his advice and support.

I wish to thank all the members of my committee for their suggestions. Dr. Steven Weintraub raised several issues during my proposal. He, in fact, gave me a three page list of corrections. After I had made the corrections, I felt that those suggestions had indeed improved the quality of this dissertation. I am grateful to him for taking the time and effort. I am also grateful to Dr. Robert Holmes, Dr. Subash Kak and Dr. Bert Boyce for their suggestions and questions during my proposal defense. I marveled at each of the issues that was raised by my committee members during my proposal defense, because this research took me through the paths that forced me to address each of those issues and on each occasion I remembered the day I proposed this topic and the questions that were asked of me then.

Any research effort has two parts. One is the advancement of science and the other the balance of life. One cannot exist without the other and my wife, Janie helped me immensely with keeping the balance, doing my share in our family. My daughters Kristi and Kristin have spent part of their life without a father, for I was always busy with either earning a living or working on this research, for most part. Their sacrifice has been immeasurable and perhaps so great a sacrifice that it has always made me wonder if the price paid for this research was too high. Perhaps it was.

My parents T.S. Vishwanathan and Sushila Vishwanathan have made very great sacrifices for me. I was very fortunate in having them as my parents. I am grateful to them for raising me and giving me the home and environment that I grew up in. Their life was again full of sacrifices, a debt I cannot repay but only enjoy and cherish and do the same for my children.

While working at the School of Veterinary Medicine, I reported to Dr. Richard Hidalgo, who was the Director of Veterinary Computing Services. Dr. Hidalgo was a constant source of encouragement. He was always understanding of my need to take classes to complete the course requirements of my Ph.D. I want to take this opportunity to express my gratitude for all the things that he has done to help me during the years that I have known him.

Dr. Helmut Schnieder employed me as an instructor in the Information Systems and Decision Sciences Department. I do appreciate the confidence he showed in me when he asked me to teach the Database Processing and Management Course for the ISDS Department.

Mr. Joe Dupard and Mr. Roy Griffin are two friends who helped me immensely in many ways and this made my life easier. I would also like to thank the many members of

my church who prayed for me. Heartfelt thanks also go to Mr. Steven Brown , Mrs. Cindy Brown, Mr. Pat Brown, Mrs. Martha Brown for their confidence in me and their support. Dr. Harold Hsu's research work in the area of wavelets and his willingness to explain concepts in a clear manner, I have always admired. I also wish to thank him for his professional encouragement. Dr. Jian Lu spent a lot of time discussing concepts that helped me improve my understanding of this subject of wavelets and I am grateful to him for his time and willingness to help. I would also like to thank Dr. Harold Stone for giving me the professional opportunity to continue this research, beyond the requirements of my Ph.D. degree.

This acknowledgment would be incomplete if I did not thank Dr. Gerald Kaiser for drilling into me concepts in wavelets and walking through my dissertation research, especially during the formative period. Dr. Jorge Aravena introduced me to wavelets from a filter theory perspective when I took a course on wavelets with him. I have since badgered him with questions on wavelets that he has always been kind enough to answer.

I would also like to take this opportunity to sincerely thank my family and friends Mr. Elias Khalaf, Mrs. Veera Watkins, Ms. Joyce Droze, Dr. Si-Qing Zheng, Dr. Doris Carver, Dr. Peter Chen, Dr. Jianhua Chen, Dr. Xian-He Sun, Dr. S. S. Iyengar, Mrs. Ramona Forbes, Mr. David Smith, Ms. Kathy Edgeworth, Mrs. Kathryn Traxler, Ms. Peggy Hatch, Mr. Malcom McNailor, Mrs. Lucille Anderson and members of the Bethany Cell Group, Mrs. Pam Dupard, Mr. Richard Bullock, Mr. George Mulkey, Mrs. Betty Mulkey, Mrs. Dorothy Mulkey, Mr. Albert Mulkey, Mr. Billy Mulkey, Dr. H. La Min, Dr. Oleg Pianykh, Dr. Kosmas Karadimitriou, Dr. Subash Babu, Mr. Tolani, Mr. Massand, Dr. Luis E. Cuellar, Dr. Pat Crawford, Dr. Richard Corstvet, Dr. Saied

Andalib, Dr. K. K. Purswani, Mr. Ashwin Iyengar, Mr. Adisak Srinakarin, Mr. Vasant Kumar, Mr. Sethu Madhava, Ms. Debora Washington, Mr. Udayan Mukherjee, Mr. Ravi Shankar, Mr. Srini Iyer, Mr. Jack Neelis, Mr. Saxon D'Aubin, Mr. Venkatraman Balasubramaniyan, and Mr. Brad Johnson. I received much consolation and moral support from them during this dissertation.

My sincere thanks and appreciation also go to Ms. Jan Day and Mrs. Janie Sharman for taking the time to read this dissertation and help me correct the errors, I did not catch. Their help comes at a crucial time, when I most need it and it will help me bring the writing of this dissertation to a close and this I have been yearning to do, for a while now.

The data for this research was provided by M.D. Anderson Cancer Research Institute and without it this research would have been difficult. Dr. John Grossman and Dr. Sidney Wallace made this data available to us. It took time and effort on their part to make this data available to us and I want to acknowledge their contribution.

It is important for me to thank you the reader for not looking at this dissertation as a work of perfection but only at that which is good in it.

From the start of this dissertation life has been a long, hard and lonely journey. This research has been possible only because of the strength and hope that comes from God, the Lord Jesus Christ. He has been my constant companion through the lonely roads of life. I owe Him everything for carrying me through impossible odds. At the conclusion of this dissertation, I can dream that the impossible is achievable because with God nothing is impossible. "Jesus looked at them and said, "With man this is impossible but not with God; all things are possible with God"" Mark 10:37. He gives to

those that believe in Him, the desires of their hearts. "Delight yourself in the Lord and he will give you the desires of your heart" Psalms 37:4. I choose His ways.

TABLE OF CONTENTS

ABSTRACT

Registration is the process of spatially aligning two objects and is normally a preprocessing step in most object recognition algorithms. Registration of images and recognition of signatures of objects in images is important for clinical and diagnostic purposes in medicine. Recognizing structure, potential targets for defense purposes and changes in the terrain, from aerial surveillance images and SAR images is the focus of extensive research and development today. Automatic Target Recognition is becoming increasingly important as the defense systems and armament technology move to use smarter munitions. Registration of images is a preprocessing step in any kind of machine vision for robots, object recognition in general, etc. Registration is also important for tuning instruments dealing with images.

Most of the available methods of registration today are operator assisted. The state of registration today is more art than science and there are no standards for measuring or validating registration procedures. This dissertation provides a viable method to automatically register images of rigid bodies. It provides a method to register CT and MRI images of the brain. It uses wavelets to determine sharp edges. Wavelets are oscillatory functions with compact support. The Wavelet Modulus Maxima technique and the Principal Component Analysis (PCA) are exploited to determine the invariant structures between images. The Modulus Maxima technique is used in isolating singularities. It also provides a mechanism to characterize the singularities in the images using Lipschitz exponents. This research provides a procedure to register images which is computationally efficient. The algorithms and techniques are general enough to be applicable to other application domains. The discussion in this dissertation includes an introduction to wavelets and time frequency analysis, results on MRI data, a discussion

on the limitations, and certain requirements for the procedure to work. This dissertation also tracks the movement of edges across scales when a wavelet algorithm is used and provides a formula for this edge movement. As part of this research a registration classification schematic was developed.

CHAPTER 1

INTRODUCTION TO ISSUES IN MEDICAL IMAGING

1.1 Introduction

This dissertation studies the registration of images with an emphasis on registering medical images. Image registration is an image processing technique of spatially aligning two similar images by rotating, translating, shearing, etc. This dissertation provides a new approach to image registration which does not require operator assistance, and is efficient in terms of computational time complexity.

1.2 Digital Image Processing

Digital image processing is now becoming widely accepted in the medical community and medical imaging of patients by physicians for treatment and diagnostic purposes is a growing trend. Digital image processing stated simplistically is the manipulation of images by a computer. At a very basic level this requires that scenes, in a naturally occurring form, be digitized by a scanner for further processing by a computer. In medical imaging several types of scanners are available for the purposes of digitizing organs of the human body. The different scanners along with the associated techniques, can image different kinds of information about body tissue, structure, etc. The images have different sizes and different gray scale digitize resolution. A further discussion on this subject is the content of the next section. In addition this chapter provides an overview of data formats used for medical images, common image processing techniques and some insight into medical data storage needs.

When a physical image is digitized, the image is stored in the form of a rectangular sampling grid. The center of each grid is the location of what is commonly referred to as the picture element called pixel. The image space is thus divided into small regions. In black and white binary images as well as in gray scale images, the brightness intensity at each pixel location is represented by an integer value called the gray level. Each pixel has an address represented by its row number and column number and a brightness level represented by the gray level. All the pixels together make up the image which can then be manipulated, stored, displayed, printed, etc. The number of pixels it takes to represent a physical picture gives a measure of the scanners quality. If a digitized picture has a smaller pixel size then the physical image is digitally represented by a larger number of pixels. The number of pixels that make up rows and columns of the rectangular grid is referred to as picture size. Another factor that determines the quality of the picture is the gray scale resolution. In most pictures the brightness at a pixel location is represented by an integer in the range 0 to 255 [256-level gray scale]. An integer in this range of 0 to 255 can be represented in eight bits using a binary number system. Since it takes eight bits to store numbers in this range, one byte of storage space is adequate. Such images are often referred to as 8-bit images. However recent trends, with the newer scanners used for medical imaging are providing better quality both in terms of size and gray scale resolution. It is not uncommon to find new scanners that provide an image size of 2048×2048 with 4096 gray levels. This constitutes 12-bit images as it takes twelve bits to store numbers in the range [0,4096]. The minimum amount of data that can be read from a disk or network channel or written to a disk or network is one byte (eight bits). However in the registers of the computer, we can use bit level information, with the help of instructions that are capable of addressing bit fields. When writing or reading from a disk, main

memory, or even for the purpose of transmission, we have to deal with bytes. So the technique of bit stuffing [Pohl84] is commonly used. For example, bit-stuffing allows programmers to store two 12-bit numbers into three bytes (twenty four bits) and read, write or transmit that information.

An image is often viewed as a two-dimensional light-intensity function, where the value or amplitude of $f(x,y)$ at spatial coordinates (x,y) gives the intensity (or brightness) of the image at that point. When an image is digitized, the digitization of the spatial coordinates (x,y) is referred to as image sampling and the amplitude digitization is called gray level quantization.

Once the images are digitized they can be subjected to a number of operations using numerical algorithms for the purpose of storage, transmission, enhancement, etc. This digital processing of images is often called digital image processing. Some of the more common image processing operations are :

- Image segmentation

  Segmentation is the process of subdividing an image into its constituent parts or objects. Each object is isolated from the rest of the image for further analysis or classification.

- Object Classification

  Object classification is the process of assigning each image or each object in an image to one of several pre-established groups (classes) that represent all the possible types of objects expected to exist in the image. This is actually a last step in an object recognition process.

- Feature Extraction

As a measurement of the value of some quantifiable property of an object. A feature is a function of one or more measurements, computed so that it quantifies some characteristic of the object.

- Edge detection

  An edge is a boundary between two regions with relatively distinct gray level properties. Edges may represent object boundaries, object shape changes, and in the case of medical images, changes in tissue density or texture within organs.

- Hidden Surface Removal

  This is done when three dimensional objects are rotated and have to be viewed on a display device, which is essentially two dimensional. The surface of the object that is not within the viewing angle is removed so that only the part that is facing the observer is displayed.

- Skeletonizing and Thinning

  Skeletonizing is the process of reducing the structural shape of a plane region to a graph via an operation called thinning. Thinning is the process of interactively deleting edge points of a region subject to the constraint that deletion of these points does not remove end points and does not break the connectedness by causing excessive erosion.

- Image Registration

  Finding a correspondence between two images. The next chapter addresses this topic extensively.

The focus of this dissertation is image registration. Specifically the focus is on medical images. Medical images provide constraints and characteristics that are specific to the field

of Medical Imaging. Hence, a brief discussion of medical imaging modalities, their resolution and storage space requirements form part of the next section.

1.3 Medical Imaging

Many of the issues in the area of medical imaging are the same as in any other area of image processing. However, there are also issues that are so specific to the area of medical imaging that have almost no impact upon general image processing algorithms. This section addresses some of the concerns that differentiate the field of medical imaging from different fields.

1.3.1 Legal Issues and Its Implication

Because of legal repercussions, compression of medical images must be lossless at this time. This reduces the compression ratios that are achievable. It affects the transmission of medical images and denies any improvements lossy images could provide. Any image processing technique must adhere to these constraints. For some of the clinical needs, clinicians and radiologists can use images with some loss of data, but clearly where one crosses this line is a subject matter of research and debate. Preliminary diagnostic procedures, simulations, training, image registration, image restoration, allow for some amount of data loss. However, compression of images must be lossless.

1.3.2 Data Formats

Medical Informatics and Radiological societies have adopted data format standards for storage, transmission and for use in medical imaging equipment that have additional constraints as compared to those that are commonly used in other areas of image processing. The data image formats that are specific to medical imaging are:

- Digital Imaging and Communications in Medicine Standard (DICOM)

- Medvission
- Network Common Data Form (NETCDF).

These are the standards that are used in medical imaging equipment, medical image transmission, medical image storage, etc. The other commonly used data formats, though to a lesser extent, in Medical Imaging are:

- Portable Graymap(PGM),
- Portable Bitmap (PBM),
- Windows Bit Map (BMP),
- Joint Photographic Experts Group Standard (JPEG),
- Graphics Interchange Format (GIF),
- Tag Image File Format (TIFF).

The Magnetic Resonance Images (MRI) and Computer Topography (CT) Images of the brain used in this dissertation were obtained in DICOM format. The DICOM images were converted to PGM format for this research. The DICOM-PGM format conversion routines were part of this dissertation research.

1.3.3 Imaging Modalities

A wide variety of image modalities are used to image patients. These modalities acquire images using very different techniques. The physics behind each of the imaging techniques is different. As a consequence of this, from a single image scene, each of these modalities digitize different information. To exemplify the issue consider a situation that a single brain is imaged with imaging modalities MR, CT and PET. The MR image would provide information on soft tissue, fatty cells, etc. and some information on bone structure, the CT image would provide the best information on denser tissue like bones, and the PET

would provide the best information with respect to blood flow, arteries and veins. More importantly, the size of the image and the gray scale resolution from each modality's scanner is different. Some of the imaging modalities in use are:

- Computed Tomography [CT]
- Magnetic Resonance Imaging [MRI]
- Positron Emission Tomography [PET]
- Single Photon Emission Computerized Tomography [SPECT]
- Magnetic Resonance Spectroscopy [MRS]
- Ultrasound [US]
- Digital Subtraction Angiography [DSA]
- Digital Fluorography [DF].

1.3.4 Medical Imaging Data

The different imaging modalities often provide varying image sizes and gray scale variation. Typically image sizes vary from images of size 256×256 to 4096×4096. The gray scale resolution varies from 8-bit pixels to 16-bit pixels. When a patient is scanned, several images of the patient's organ are obtained, usually a few millimeters apart. For example when MRI images are obtained each exam would result in obtaining about 50 images. Data from a study done a few years ago is presented in the table below and provides a measure of the volume of data and gray scale variation with imaging modality. This data is obtained from Rabbani[Rabbani91]. Because of continuing improvements in scanner technology most new scanners nowadays are providing image sizes which are larger and with gray levels ranging from 0 to 4096 (16-bit images). Usually patients in large cancer research centers, are examined and imaged once a month. When the data storage requirements are

all summed, the storage requirements are now in Terabytes and growing. Storing and transmitting of this data for tele-radiology purposes is a major new issue.

Table 1-1 : Data generated for different medical imaging modalities.

Modality	Image Dimension	Gray Level	Average Number of images per exam	Average Mbytes per exam
CT	512 x 512	12	30	16
MRI	256 x 256	12	50	6.6
DSA	1024 x 1024	8	20	20
Digital Fluorography	1024 x 1024	8	15	15
Ultrasound	512 x 512	6	36	10
SPECT	128 x 128	8 or 16	50	0.8 or 1.6
PET	128 x 128	16	62	2
Digitized Film	2048 x 2048	16	4	32

Image analysis and image processing algorithms such as image registration, image segmentation, image compression, image classification, etc., have to deal with multi-modal as well as uni-modal images of varying size and resolution. While some human organs are almost rigid, many other organs have huge shape transformations depending on factors like the posture of the body, food and liquid ingestion, etc. Image analysis and image restoration issues become quite complex in this environment.

1.4 Conclusion

The focus of this dissertation research is only registration of medical images. Registration is a process of mapping one image onto another. Images are also registered for a variety of other needs ranging from object recognition for military purposes, automatic navigation of robots through terrain, recognition of land terrain, to determine changes in the geography of real estate, rectification of display distortions, map and cartographic projections, automatic adjustment of instruments, etc. The medical imaging registration is usually done to detect regional growth such as detecting bone assimilation in the case of bone grafting, the growth of cancerous cells, for general diagnostic purposes, and for improved set compression purposes as proposed by Karadimitriou and Tyler [Karadimitriou97]. Medical Image matching is another rapidly evolving field of interest, with its own unique operational conditions.

The next chapter addresses the issues related to image registration. A literature survey of techniques used in the registration of images, as well as a classification of these techniques, is also included as part of this chapter.

CHAPTER 2

IMAGE REGISTRATION - TECHNIQUES, ANALYSIS AND A LITERATURE REVIEW

2.1 Introduction

Why should we register objects? What purpose does it serve to register images? At the outset, this chapter attempts to answer these questions at the outset. The remainder of this chapter is a review of the literature on medical image registration. The emphasis has been to focus on the techniques for of image registration. Since this dissertation is devoted to the development of a technique of registration where the control points (homologous points) are determined using wavelets, the medical aspects of the treatment, the radiation doses needed, or other purely clinical aspects are not be emphasized. Though, some of the content is devoted to external markers, this is also not the focus of this literature survey. These sections have been included only for completeness and to demonstrate that we are aware of these techniques to obtain control points. After reviewing the literature on the techniques to register objects, this chapter provides a schematic diagram that lays out a basis by which registration techniques can be classified. This schematic in Figure 2-1 is a direct outcome of the literature survey in this field. This chapter also presents a model for registering images and shows how correlation coefficient is related to a mean square difference or euclidean distance between the images being registered. Each image in this case is viewed as a normalized vector with mean zero. As the mean square distance decreases the correlation coefficient increases. Intuitively it is easy to see that as images are better spatially aligned the mean square distance between them decreases and the correlation coefficient increases.

2.2 The need for registration

Registration of two images A and B means the determination of translation and rotation parameters, which when applied to say image B, would spatially match the other image, i.e. image A. For rigid bodies there is no distortion but the transformations are affine. Registration in a broader perspective includes the need to shear or in some way distort one of the objects to provide a match to the second object or template. This is usually done to rectify distortions.

There is a great need to register objects in almost all areas of image processing including the area of medical imaging. The human eye constantly registers objects for the purpose of object recognition, object classification, and this registration is superior to any other known technique. For this reason, all of the commercial image registration algorithms available for registering medical images, require operator assistance. Research in the area of automatic registration is still in its infancy. There are no commercial automatic registration tools to register medical images in the market today. However, the army and navy are exploring the area of image registration and research is moving toward automatic image registration. This dissertation research deals with registration of Magnetic Resonance brain images. These can essentially be treated as registration of rigid bodies. To provide a deeper appreciation for registration, the next few paragraphs detail some areas where there is a great need for image registration:

- In machine vision, registration of scene images is a fundamental operation. Useful data is derived only after registration and fusing of information obtained from multiple sensors. This is especially true for robots attempting underwater navigation, navigation through unknown terrain. Registration is almost always necessary for robots to

recognize objects. For example, a robot in an automated factory would first register as a step in the recognition of an object then classify and place it at the appropriate bin.

- For aerial images, the matching operations are needed for a variety of applications. For example, change detection, map updating, aerial surveillance, and target recognition warrants its use during war time. Registration has to be done many times if there is a changing terrain. In aerial images, the images are obtained from an altitude higher than the tallest structure in the scene. The position of the sun when the image is obtained, the cloud cover, etc., are issues that have to be considered when registering aerial images.
- In medical imaging, different modalities provide additional complimentary information. Single Photon Emission Computed Tomography (SPECT), Positron Emission Tomography (PET), and Magnetic Resonance Spectroscopy (MRS) provide functional information, but delineate anatomy poorly. Magnetic Resonance Imaging (MRI), Ultrasound (US) and X-ray imaging and Computed Tomography (CT) depict aspects of anatomy, but provide little functional information. Computed Tomography (CT) provides better information on denser tissue like bone, but Magnetic Resonance Images (MRI) provides more useful information on soft tissue, for example, fatty tissues, cancerous regions, etc. In patient diagnosis, there are many situations that warrant the need for additional information from the different imaging modalities and then integrate this into a useful form. This requires images to be registered before integration (fusion). An example of an application area for multimodality would be radiotherapy and nuclear medicine. In radiation therapy planning, a CT scan is needed for dose distribution calculations, while the contours of the target lesions are best outlined with MRI. Determination of anatomical location of dysfunctional areas and

studies on functional-structural relationships are facilitated by integration of functional and morphological information. This integration is achieved through image registration and is another example application.

Registration of monomodal images in medical imaging is useful in applications where atlas data are used, for example in studies over a patient population. Registration of monomodal images is essential in monitoring a patient's progress when treatment is administered over a period of time. This is usually the case with patients undergoing cancer treatment. Such studies are also done to closely monitor how well a bone graft is taking.

Registration, according to Lisa Gottesfeld Brown [Brown92], is often necessary for:

1. integrating information taken from different sensors,
2. finding changes in images taken at different times under different conditions,
3. inferring three-dimensional information from images in which either the camera or the objects in the scene have moved, and
4. model-based object recognition.

2.3 Techniques Used in Registration

Many classes of methods have been developed for image registration. Classification of these methods can be done on several basis [Barnea72][Elsen93][Elsen91]. Based on the literature survey of the techniques used in image registration, the following modified basis for classification is offered :

2.3.1 Level of Abstraction

Registration methods can be categorized by the level of abstraction. The description hierarchy at which the matching is performed could be at the pixel level, the simplest case, up to matching based on high level object derived descriptions, like roads, buildings,

arteries, anatomical landmarks, etc. At the pixel level the available descriptors are pixel properties such as intensity, color, gray level, etc. The matching at the pixel level can be based on some form of area correlation [Brown92][Herbin89]. The intensities may change based on imaging conditions and multimodal registration may require gray level normalization. At the other extreme, matching is done based on derived object descriptors that are context sensitive [Brooks89]. For example, recognizing semantic structures like roads, buildings, reservoir, lakes, major arteries, certain anatomical structures that could be used by matching these structures and their interrelationships. The higher the level of descriptions at which the matching is attempted, the more likely the descriptions are invariant to changes in the imaging conditions [Medioni84]. However, the harder part is deriving image descriptors that are consistent across the available set of images on which digital registration is attempted. Moreover, to some extent, some level of abstraction is helpful but that has to be based on the context of the images. This dissertation makes use of shapes of structures that stay invariant across the images being registered.

2.3.2 Matching based on the number of dimensions

Matching can be performed at the one dimensional, two dimensional or three dimensional level. When a one dimensional level is used, the image is essentially treated as a one dimensional signal. This reduces to finding the ideal delay. In filter theory parlance, it translates into processing the image signal through a filter process that has a linear phase thus preserving the waveform but introducing ideal delays. Without some criteria that is able to recover from the local minima, this process may not find the global minima. This matching process may have to deal with such issues. Speech signals are usually processed as one dimensional signals. Two dimensional methods are based on several criteria. For most clinical studies that require multimodal registration, the assumption is made that the

images were generated in the same plane relative to the patient. Depending on the objective function being optimized in the registration process in a particular study, like that of compressing sets of images, this assumption need not be made. However in most cases, it is assumed that the images being generated are similar. Objects of different types are not registered in the normal case. In medical imaging it is assumed that the image slices obtained using different modalities, or the same modality, are views of the same scene or organ. In aerial surveillance when incoming missiles or enemy war planes have to be identified, each satellite provides a two dimensional picture of the same object but from different angles. To do a target recognition, the images have to be first registered and then fused. Three dimensional methods in the area of medical imaging, do not consider tomographic images as a set of individual slices but consider them as a volumetric data set that can be registered with another three dimensional volumetric data set. However, it must be pointed out that this volume is constructed from two dimensional slices. This necessitates extrapolating between the slices and thus creation of data. This type of registration is useful in clinical surgery preparation and other diagnostic needs. Since storage and archival of medical data cannot afford any distortion or creation, however accurate the process of extrapolation, legal requirements do not, at this time, permit three dimensional registration for compression purposes [Clarysse91][Elsen91][Faber91]. In this dissertation, registration is performed at the two dimensional level. The ideas, algorithms and techniques can be extended to the three dimensional level.

2.3.3 Origin of Image Properties

Registration algorithms may either use patient related intrinsic image properties or extrinsic image properties that are added to the patients. Intrinsic properties are pixel intensities, anatomical landmark points, geometrical features like cortex, ventricles, etc.

[Rousseau92]. Materials like radio pharmaceuticals, contrast media for angiograms, etc., administered to patients to enhance the contrast, are considered intrinsic properties. Examples of extrinsic properties are head frames, skin markers, etc. In this dissertation only intrinsic properties are considered.

In stereotactic procedures, where targets deep within the brain are approached through a small hole in the skull or by a narrow beam of high energy ionizing radiation, a reference frame is attached to the patients head by screws fixed to shallow drill holes in the skull. Provisions are made to ensure accurate replication of the frames. These provide commonly used extrinsic markers. Some frames can be attached to the head in a more patient-friendly manner [Sharman96]. Plastic mold or dental mold are also widely used as extrinsic markers. These markers are then used as control points. The Wavelet Modulus Maxima approach used in this dissertation could detect both extrinsic or intrinsic markers as the edge effects that produce points of singularity [Sharman97.1][Sharman97.2].

2.3.4 Elasticity of Transforms

Image matching transforms can be classified as rigid affine, projective or curved. Each of these categories indicate the degree of elasticity and shearing used in the transformation. A transform is called rigid if the distance between any two points in the first image is preserved when these two points are mapped in the second image. Rigid body transformations can be decomposed into translation, rotation and mirror reflection. Let (x, y) represent a point in one of the two images being registered and (x', y') be the corresponding point in the second image, which will ultimately be matched with the point (x, y) in the first image. Mathematically the transformation that relates the two points (x, y) and (x', y') in the two images being registered is :

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \qquad \text{Eq. 2-1}$$

where θ denotes the rotation angle and $\begin{bmatrix} t_x \\ t_y \end{bmatrix}$ denotes the translation vector[Elsen93]. The two points (x, y) and (x', y') in the two images being registered are referred to as homologous points in the literature on this subject. For this reason the same nomenclature has been adopted in this dissertation.

Since the images that are used in this dissertation, are brain images obtained through MRI and CT scans, this research deals with rigid body transformations.

Affine transformations are transformations where any straight line in the first image is mapped onto a straight line in the second image, while parallelism is preserved. Affine transformations in two dimensions may be represented as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \qquad \text{Eq. 2-2}$$

where $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is a real-valued matrix [Dougherty87].

A projective transformation maps a straight line in the first image onto a straight line in the second image but parallelism between the lines is not preserved. Curved transformations are also projective transformations but in these tranaformations a straight line is mapped into a curved line.

Some transformations call for uniform scaling in the x, y and z directions. Since organs undergo shape changes, non-uniform scaling transformations may need to be considered. The plasticity property is very difficult to model, even though some of the organs undergo

such distortions [Wolberg90]. Registration, in such cases need very careful and extensive human intervention from an expert.

2.3.5 User Interaction

Most current registration processes are semi-automatic. An expert user identifies points on each image that are homologous. Registration is then done manually based on matching the homologous points. In some cases, the user identifies points and other structures like lines, organs, blood vessels, etc., and then registration is done based on these derived abstractions as well as the points. State-of-the-art commercially available registration software for medical imaging is semi-automatic. Research in fully automated registration techniques is in its infancy.

In a fully automated registration process, the points or structures are automatically detected by an image processing program. It is for this reason that wavelet based registration should make a significant contribution to the fields of medical imaging and registration. To keep within the scope of this dissertation, only rigid body transformations have been considered. MRI and CT scans of the brain can be considered rigid bodies. Within these rigid bodies, search algorithms can be used. but, exhaustive search methods are computationally expensive. In this research, wavelet based techniques are used to detect singularities in known geometrical positions and to locate homologous points and then register the images. These wavelet based techniques are not considered to be computationally expensive [Bruce96]. The method considered in this dissertation provides a way to further reduce this time complexity.

2.3.6 Other considerations

A point-to-point match method for registration can be done by using a Singular Value Decomposition (SVD) technique to find a least squares fit. Alpert, Kennedy and Correia

[Alpert90], in their study use SVD to register PET and CT data sets. In this study accurate anatomical localization was needed and multimodal registration was necessary to treat tumors. They knew that a tumor involved nodes that had increased glucose metabolism and this was used to manually obtain the homologous points. Evans et. al [Evans89] used the SVD technique to register PET and MRI of brain images, while Hill, et. al [Hill91] used SVD to register point-to-point CT and MRI skull base images successfully. SVD is often referred to as Principal Component Analysis or the K-L Transform. The SVD of the covariance matrix of a cluster of points chosen automatically or with user help, is used to define the principal energy axis. The SVD also gives the variance in directions perpendicular to the energy axis line. To elaborate, the general idea of Principal Component Analysis, is to use the centroid of the points and find the eigenvectors of the covariance matrix. The points used in the calculations, are those that are identified by a radiologist. In this dissertation we use points obtained by using the Wavelet Modulus Maxima. These eigenvectors of the covariance matrix describe the axis of minimum least square distances for the given set of points. This yields an orthogonal object reference system. The same process is applied to the second image, that is to be registered with the first image. This information is then used to derive the parameters of the transform. This transformation consists only of rotations and translations [Toennies88]. Another advantage is that the eigenvalues associated with the eigenvectors of the covariance matrix represent magnifications, and may be used to specify scaling parameters for applications that involve changes in magnification. Meyer et. al [Meyer95] not only used points but line and plane descriptors in the registration process to find the Least Mean Square error during the optimization.

Another frequently used method in medical image registration is based on matching moments. Points and other landmarks are used in this process. The moment of inertia is computed and from that the principal axis of the object is calculated. The extent of rotation is computed based upon aligning the principal axis of inertia and then the translation parameters are computed based on the centroid of the homologous points and landmarks.

Correlation based methods map regions of pixels on a best fit basis. This is widely used in image registration. Most of these techniques are semi-automatic. The hypothesis of this research is that using wavelets one can detect points of singularities in the images to be registered. This process does not require any user interaction. The images being registered have known landmarks that can be used for registration. These are located in certain regions of the image and the Wavelet Modulus Maxima can help identify these features. The points identified by the Wavelet Modulus Maxima form structures that are shape invariant. We use principal components of the shapes in the registration process. In the case of brain images the shape of the skull is invariant. This observation allows us to use the SVD or moment of inertia technique to register the objects. The optimization function in most registration algorithms is a mean square error reduction among the homologous points. However, if the purpose of the registration is set image compression, then the optimizing function should be the difference image minimization. However a quantitative measure is the correlation coefficient and the goal is to maximize it by rotation and translation of the images.

2.4 Mathematical Model for Registration

Let $f_1(x,y)$ and $f_2(x,y)$ represent the two images that are to be registered.

Let

- $T\begin{bmatrix} x \\ y \end{bmatrix}$ represent the translation operator or transform. T^{-1} is the inverse of T in the sense $T^{-1}\begin{bmatrix} x \\ y \end{bmatrix} = T\begin{bmatrix} -x \\ -y \end{bmatrix}$. Since this is a linear operation the inverse exists.
- $R(\theta)$ represent the rotation operator or transform. $R^{-1}(\theta)$ is the inverse in the sense $R^{-1}(\theta) = R(-\theta)$.
- M be the scaling or magnification operator. M^{-1} is the inverse of M providing for shrinking of the image.
- G_L be the gray level adjustment that may be needed.

Since we are considering rigid body registration in this study, plasticity is not considered. The correlation coefficient, r, between two images is defined as

$$r = \frac{\sum_{x=0}^{x=IMAGESIZE-1}\sum_{y=0}^{y=IMAGESIZE-1}\left[f_1(x,y)-\overline{f_1(x,y)}\right]\left[f_2(x,y)-\overline{f_2(x,y)}\right]}{\left\{\sum_{x=0}^{x=IMAGESIZE-1}\sum_{y=0}^{y=IMAGESIZE-1}\left[f_1(x,y)-\overline{f_1(x,y)}\right]^2 \sum_{x=0}^{x=IMAGESIZE-1}\sum_{y=0}^{y=IMAGESIZE-1}\left[f_2(x,y)-\overline{f_2(x,y)}\right]^2\right\}^{1/2}} \quad \text{Eq.2-3}$$

Let $f_3(x,y)$ represent the image after registration obtained by translating, rotating, magnifying (or shrinking) and adjusting the gray levels in image represented by $f_2(x,y)$.

$$f_3(x,y) = R(\theta)T\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}MG_L f_2(x,y) \quad \text{Eq. 2-4}$$

If we consider each of the images $f_1(x,y)$ and $f_3(x,y)$ to be normalized vectors with zero mean we have the relationships

$$\sum_{x,y} f_1^2(x,y) = 1 \quad \text{Eq. 2-5}$$

$$\sum_{x,y} f_3^2(x,y) = 1 \qquad \text{Eq. 2-6}$$

$$\sum_{x,y} f_1(x,y) = 0 \text{ and} \qquad \text{Eq. 2-7}$$

$$\sum_{x,y} f_3(x,y) = 0 \qquad \text{Eq. 2-8}$$

We can now establish the relationship between the mean squared difference (D^2) between the images and the correlation coefficient (r). The mean squared difference (D^2) is a measure Euclidean distance between the images and the correlation coefficient is the cosine of the angle between the images.

$$D^2 = \left(f_1(x,y) - f_2(x,y)\right)^2 = 2(1-r) \qquad \text{Eq. 2-9}$$

The objective of the registration is to minimize D^2 which is the same as maximizing the correlation coefficient (r).

2.5 Conclusion

This chapter presents some of the more commonly used methods in registration. There are several variants of these basic techniques that have been used for the purposes of registration. Based on the literature survey done as part of the research for this dissertation, the schematic in Figure 2-1 was developed to represent the classification of registration techniques. This research uses a novel and automated approach for locating the homologous point pair. Moreover, once the homologous point pairs are obtained, the transform parameters are computed using a new and different correlation measure. Most of the research efforts in registration provide no quantitative and conclusive way of determining the accuracy of the registration process. In this work the objective is to minimize the difference image by maximizing the correlation coefficient between the

images being registered. There is a strong likelihood that this process will be totally automated. These are major motivations for this research project. Medical images may be stored using wavelets in the future. The use of wavelets is being widely touted by researchers in Medical Imaging (hospitals and commercial establishments). If wavelets become the accepted industry standard for Medical Imaging, the results of this project will probably be used as a front-end. The results of this project can also be implemented in registration of Landsat images.

To comprehend wavelets and this application, it is imperative that one understands Fourier Transforms and issues related to time-frequency analysis. The next chapter is devoted to explaining concepts of aliasing, frequency domain analysis and wavelets in general.

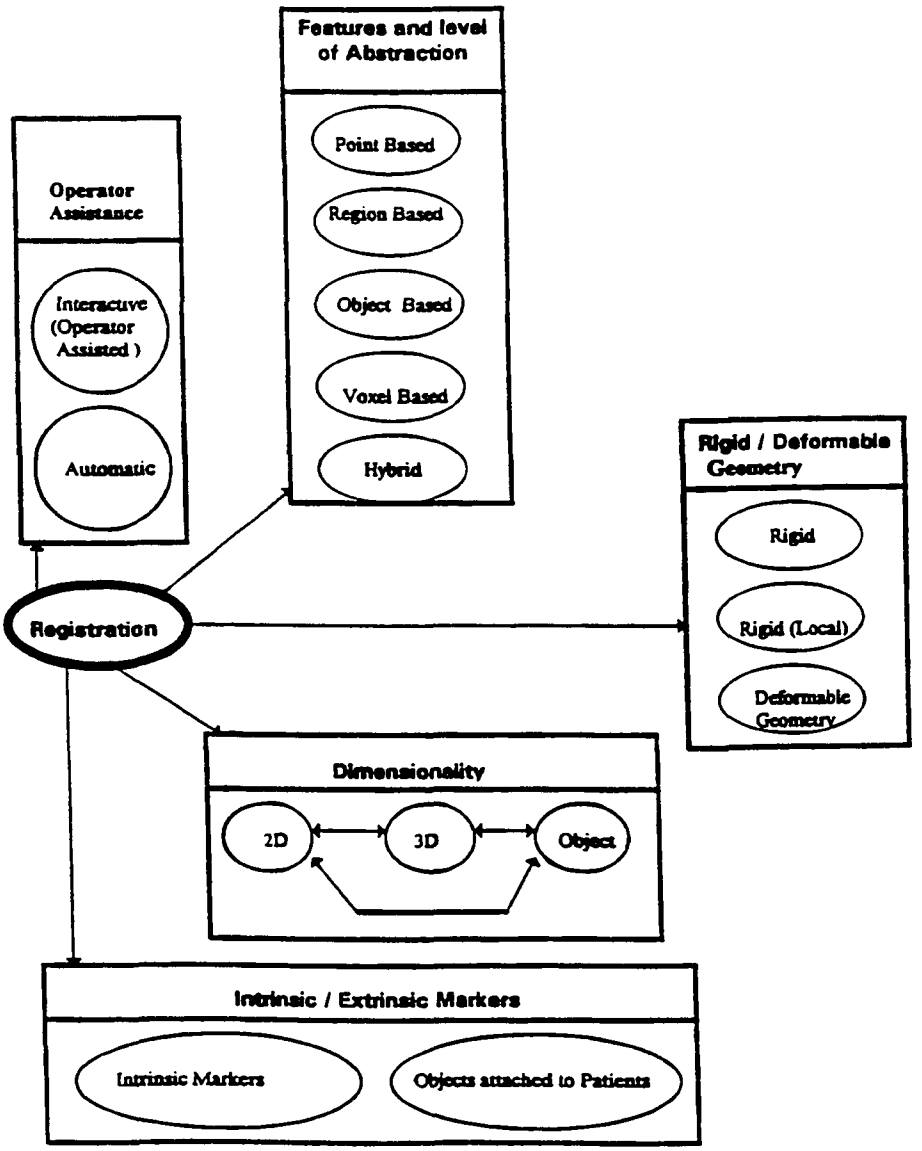


Figure 2-1 : Classification of Techniques to register images

CHAPTER 3

TIME FREQUENCY ANALYSIS AND THE WAVELET TRANSFORM

Wavelets are used in this dissertation. Wavelets are new functions that are used to represent signals. They are alternatives to Fourier transforms. Wavelets are short waves instead of long waves. Figure 3-1 shows a picture of sinusoidal wave and Figure 3-2 shows a function that decays rapidly. Figure 3-3 shows the Mexican Hat Wavelet. The successful use of wavelets depends upon the actual signal and the application for which the analysis is done.

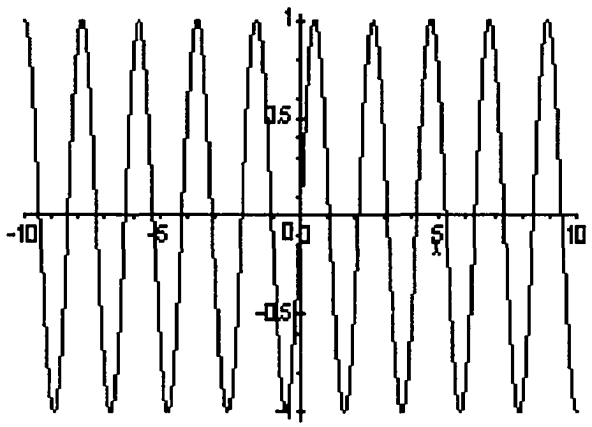


Figure 3-1 : Sine wave (does not decay)

To explain wavelets and its impact, we use Fourier techniques - the connection between time and frequency, as this idea remains the center of signal processing. Wavelets are alternatives rather than replacements for stationary signals. However for non-stationary signals Wavelet Transforms provide a better analysis than Fourier Transforms. This new

Wavelet Transform is much more local. Instead of transforming the time description into the pure frequency description, this new method finds a good compromise - a time frequency description [Strang96]. Wavelets are ideal for analyzing non-stationary signals because of it localization property.

A function that decays to zero

Figure 3-2 : A function that decays and also does not oscillate

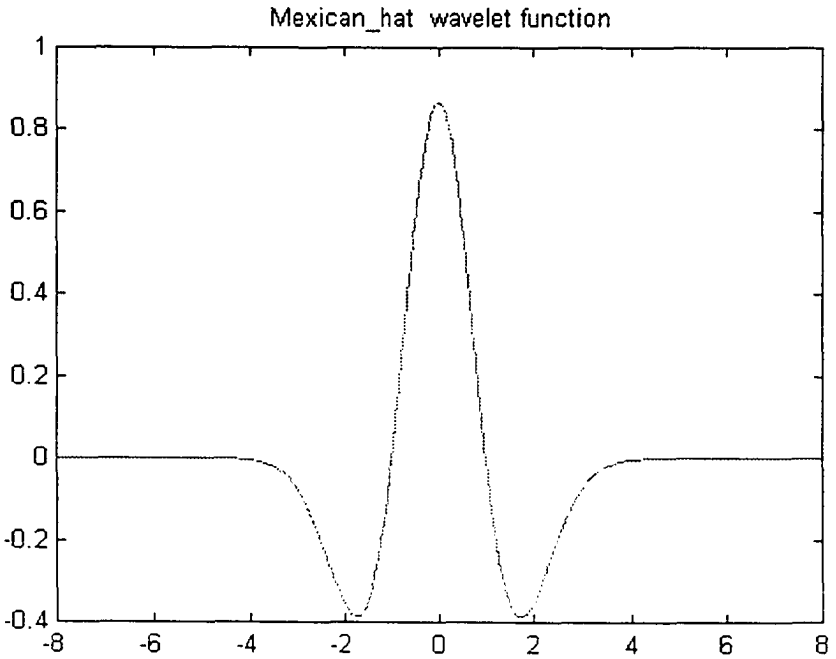


Figure 3-3 : Mexican Hat Wavelet (Function decays as well as oscillates)

3.1 Introduction

A background for wavelet research includes background and knowledge of Fourier Transforms and time frequency issues. This chapter is devoted to Fourier transforms and those time frequency issues that are needed as an introduction to wavelets and the Wavelet Transforms. An introduction to some of the other transforms, for example the Walsh Transform, Wigner Distribution (WD) and the Windowed Fourier Transform (WFT) or the Short Time Fourier Transform (STFT) are also included because they have a role in this background material. This discussion focuses on the differences in the time frequency plane, when these transforms are used.

3.2 Definitions

Fourier Transforms provide a spectrum showing the frequency content of the signal. To explain some of the issues involved in signal processing and in frequency domains, the definition of the important terms are provided here.

- Hilbert space is a space that has an inner product. An inner product is defined as

$$\langle f , g \rangle = \int f(t)\, g(t)\, dt \qquad \text{Eq. 3-1}$$

where if $L^2(R)$ denotes the Hilbert space of measurable functions, then $f \in L^2(R)$ and $g \in L^2(R)$.

Gerald Kaiser provides a more comprehensive notion of a Hilbert space. A Hilbert space H is any vector space (for example $l^2(Z), l^2(N)$ and $L^2(R)$) with an inner product satisfying [Kaiser94]:

1. Positivity : $||f|| \rangle 0$ for all $f \in H$ with $f \neq 0$, and $||0|| = 0$
2. Hermiticity : $\langle \overline{f, g} \rangle = \langle g, f \rangle \forall f, g \in H$

3. Linearity : $< f, cg + h >= c < f, g > + < f, h > \forall f, g, h \in H, c \in C$

A Hilbert space is the space of all functions $f(x)$ with finite energy. All speech signals and signals from images are finite energy signals. Let $L^p(R)$ also denote the Hilbert space of measurable functions, then

$$\int_{-\infty}^{\infty} |f(x)|^p \, dx < +\infty \qquad \text{Eq. 3-1}$$

Let $L^2(0,2\pi)$ denote the collection of all measurable functions f defined on an interval $(0,2\pi)$ with

$$\int_0^{2\pi} |f(x)|^2 dx < \infty \qquad \text{Eq. 3-2}$$

- The classical norm $f(x) \in L^2(R)$ is given by

$$\|f\|^2 = \int_{-\infty}^{\infty} |f(x)|^2 \, dx \qquad \text{E q. 3-3}$$

- The convolution of two functions $f(x) \in L^2(R)$ and $g(x) \in L^2(R)$ is defined as

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(u) g(x-u) du \qquad \text{Eq. 3-4}$$

where $*$ denotes the convolution operator. In the discrete case using a digital filter with filter coefficients $h[n]$, the convolution is

$$y[n] = \sum_{k=-\infty}^{\infty} x[k] h[n-k] \qquad \text{Eq. 3-5}$$

with output $y[n]$, and input $x[n]$ representing some signal and filter impulse responses at discrete times $t = nT$ and where T is assumed to be 1 [Oppenheim89].

In an analogous manner, the two dimensional convolution of a signal $f(x,y) \in L^2(R^2)$ with a filter function $h(x,y) \in L^2(R^2)$ is given by

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\tau_1,\tau_2)h(x-\tau_1,y-\tau_2)d\tau_1 d\tau_2 \qquad \text{Eq. 3-6}$$

The discrete convolution in two dimensions of two functions :

1. a signal $f(x,y)$ represented in the form of a discrete array of size $A \times B$ and
2. a set of filter coefficients $h(x,y)$ represented in a discrete array of size $C \times D$ is defined by the relation:

$$f(x,y)*h(x,y) = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)h(x-m,y-n) \qquad \text{Eq. 3-7}$$

where $x = 0,1,2,3,\ldots M-1$ and $y = 0,1,2,3,\ldots N-1$ and where $*$ denotes the convolution operator. The $M \times N$ array represents one period of the discrete convolution in two dimensions, if M and N are chosen such that $M \geq A+C-1$ and $N \geq B+D-1$.

- The Fourier Transform $F(w)$ of a continuous function $f(x)$ is defined as

$$F(w) = \int_{-\infty}^{\infty} f(x)e^{-iwx}dx \qquad \text{Eq. 3-8}$$

where the inverse Fourier Transform is given by

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(w)e^{iwx}dw \qquad \text{Eq. 3-9}$$

In the discrete case we have the Discrete Fourier Transform given by:

$$F(k) = \frac{1}{N}\sum_{n=0}^{N-1} f[k]e^{\frac{-2\pi kn}{N}} \quad \text{where } k \in [0, N-1] \qquad \text{Eq. 3-10}$$

and the inverse Discrete Fourier Transform is given by

$$f(n) = \sum_{k=0}^{N-1} F(k) e^{\frac{2\pi i k n}{N}} \quad \text{where } n \in [0, N-1] \qquad \text{Eq. 3-11}$$

The Fast Fourier Transform (FFT) is functionally the same as the Discrete Fourier Transform (DFT) and they yield the same results. However the FFT has a time complexity of $O(n \log n)$ and is profoundly faster than the DFT which has time complexity of the order $O(n^2)$ [Bracewell86].

The Fourier transform of a signal $x[n]$ is $X(e^{iw})$. The signal is in the time domain and $X(e^{iw})$ is in the frequency domain. The time variable n is discrete and the frequency variable w is continuous. $X(e^{iw})$ is 2π periodic.

- Many sequences can be represented by a Fourier integral of the form

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{iw}) e^{inw} dw \qquad \text{Eq. 3-12}$$

where $X(e^{iw})$ in discrete-time (signal processing notation) is:

$$X(e^{iw}) = \sum_{n=-\infty}^{\infty} x[n] e^{-iwn} \qquad \text{Eq. 3-13}$$

In general, the Fourier Transform is a complex-valued function of w. In "rectangular" form the Fourier Transform can be expressed as :

$$X(e^{iw}) = X_R(e^{iw}) + iX_I(e^{iw}) \qquad \text{Eq. 3-14}$$

In "polar" form, the Fourier Transform can be expressed in terms of the magnitude $|X(e^{iw})|$ and phase $\angle X(e^{iw})$ as

$$X(e^{iw}) = |X(e^{iw})| e^{i\angle X(e^{iw})} \qquad \text{Eq. 3-15}$$

The reason for providing this expression in "polar" form is to use the phase in determining the direction of the gradient in a two dimensional vector field.

- Another point is that, by Parseval's energy equality, for finite energy signals, the energy is the same in the time domain and the frequency domain

$$\text{Total Energy} = \int_{-\infty}^{\infty} |f(t)|^2 \, dt = \int_{-\infty}^{\infty} |F(w)|^2 \, dw \qquad \text{Eq. 3-16}$$

For any function $f(x)$, let $f_a(x)$ denote the dilation of $f(x)$ by a scale factor of a. Then

$$f_a(x) = \frac{1}{|a|} f(\frac{x}{a}) \qquad \text{Eq. 3-17}$$

- For a two dimensional case, let $L^2(R^2)$ be the Hilbert space of measurable, square-integrable functions. The classical norm of $f(x,y) \in L^2(R^2)$ is given by

$$\|f\|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x,y)|^2 \, dxdy \qquad \text{Eq. 3-18}$$

- The two dimensional Fourier Transform is given by the formula [Weaver83]

$$F(w_x, w_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) e^{-i(w_x x + w_y y)} \, dxdy \qquad \text{Eq. 3-19}$$

- The convolution theorem states that the convolution in the time domain is equivalent to multiplication in the frequency domain. Similarly convolution in the frequency domain is equivalent to multiplication in the time domain. Using the functions that were used to define the one and two dimensional convolutions, we can represent the relationships as:

For the one dimensional case

$$f(x) * h(x) \Leftrightarrow F(w) \bullet H(w) \qquad \text{Eq. 3-20}$$

$$F(w) * H(w) \Leftrightarrow f(x) \bullet h(x) \qquad \text{Eq. 3-21}$$

where $F(w)$ and $H(w)$ are the Fourier Transforms of the functions $f(x) \in L^2(R)$ and $h(x) \in L^2(R)$ respectively and w is the variable denoting the frequency.

For the two dimensional case:

$$f(x,y)*h(x,y) \Leftrightarrow F(w_x,w_y) \bullet H(w_x,w_y) \qquad \text{Eq. 3-22}$$

$$F(w_x,w_y)*H(w_x,w_y) \Leftrightarrow f(x,y) \bullet h(x,y) \qquad \text{Eq. 3-23}$$

where w_x and w_y represent frequencies in the two dimensions x and y. $F(w_x,w_y)$ and $H(w_x,w_y)$ are the Fourier Transform of the functions $f(x,y) \in L^2(R^2)$ and $h(x,y) \in L^2(R^2)$ respectively. The * denotes the convolution operator and • denotes the multiplication operator.

- For any function of $f(x,y) \in L^2(R^2)$, let $f_a(x,y)$ denote $f(x,y)$ dilated by a factor a. Then

$$f_a(x,y) = \frac{1}{|a|^2} f(\frac{x}{a},\frac{y}{a}) \qquad \text{Eq. 3-24}$$

A translation operation applied to a function $\psi(t)$ in combination with a scaling action, this composite operation, is referred to as an affine operation. The affine operation simultaneously scales and translates the independent variable. Mathematically (assuming real numbers) the affine operation maps the original variable x into a new independent x' by the formula

$$x' = \frac{x-b}{a} \qquad \text{Eq. 3-25}$$

where a is the scale parameter and b is the translation. It should be noted that affine operations are not commutative.

3.3 Digital Signal Processing

Suppose a signal $f(t)$ is represented by N sample points separated by Δt. The total time interval over which the signal is sampled is:

$$T = N\Delta t \qquad \text{Eq. 3-26}$$

where T is the truncation window.

3.3.1 Sampling

Continuous signals sampled at a frequency of twice the Nyquist frequency can be reconstructed perfectly by the Whittaker-Shannon sampling theorem. For a digital computer to process any signal it has to be sampled. When we sample a signal $x_c(t)$, we normally do so at a uniform sampling rate for a fixed period of time. The process of sampling can be viewed as using an impulse train modulator followed by the conversion of these values to a sequence. The difference between the sampled signal values $x_s(t)$ and the sequence $x[n]$, is that $x_s(t)$ is a continuous signal (specifically an impulse train) that is zero except at integer multiples of T. The sequence $x[n]$, on the other hand, is indexed by the integer variable n, which, in effect, introduces time normalization. This means $x[n]$ contains no explicit information about the sampling rate. In other words, samples of the continuous signal $x_c(t)$ are represented by the finite numbers $x[n]$ rather than as the area of impulses in $x_s(t)$. When sampled speech signals or signals obtained from any process, even image signals are examined, they are basically a uniformly sampled sequence $x[n]$. The Fourier Transform consists of periodically repeated values of the Fourier Transform of the continuous signal $x_c(t)$. The periodically repeating copies of the Fourier Transform of $x_s(t)$ are shifted by the sampling frequency $\frac{1}{T}$ and then

superimposed to produce the Fourier Transform of the sequence $x[n]$. Figure 3-4 shows the continuous-time Fourier spectrum of a band limited signal and that of a sampling function. It also shows the spectrum after the two signals are convoluted with no aliasing and with aliasing.

Since a signal can be sampled with a finite number of points, this sampling process truncates the signal by ignoring any value of the signal that is outside the truncation window. This has the same result as setting all values of the signal to zero outside the window. Since the sampled signal consists of N independent measurements, it is reasonable to compute using a total of N points on the spectrum. Since we assume $f(t)$ is sampled with spacing Δt, its spectrum, $F(s)$ is periodic with period $1/\Delta t$. If we consider N evenly spaced sample points on the spectrum $F(s)$ of $f(t)$ and consider one cycle of this periodic spectrum centered about the origin; then N points computed by either DFT or FFT on $F(s)$ is over the range $-\frac{1}{2\Delta t} \le s \le \frac{1}{2\Delta t}$. If the N points in the Frequency spectrum are separated by an interval Δs, then the entire period is $N\Delta s$. Hence

$$N\Delta s = \frac{1}{\Delta t} \qquad \text{Eq. 3-27}$$

implies

$$\Delta s = \frac{1}{N\Delta t} = \frac{1}{T} \qquad \text{Eq. 3-28}$$

If we consider a band limited signal with s_m as its highest frequency, then $s_m = \frac{1}{2\Delta t}$. The maximum frequency that can be computed is inversely related to the time domain sample spacing.

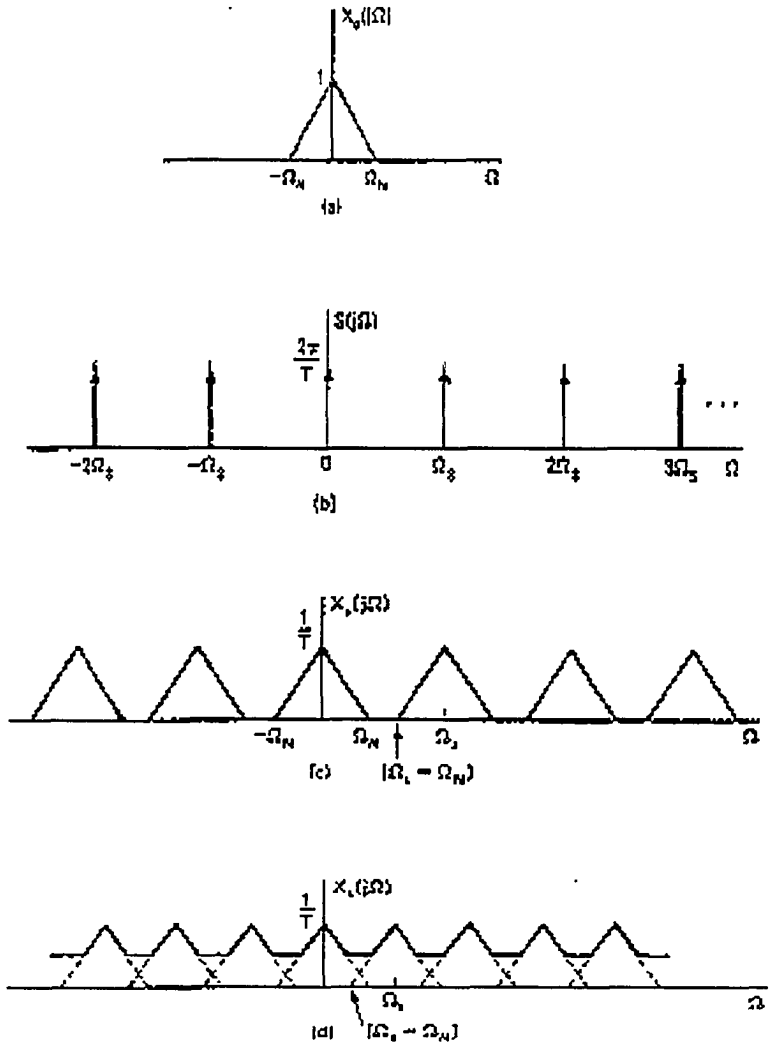


Figure 3-4 : This figure shows (a) Fourier spectrum of a band limited signal (b) Fourier spectrum of a sampling function (c) Spectrum of the sampled signal without aliasing and (d) the spectrum of the sampled signal with aliasing [Oppenheim89].

The sample spacing in one domain dictates (or is dictated by) the truncation width in the other domain. If we compute high-frequency components of the spectrum, then we must finely sample the time domain. High resolution in the spectrum (small Δs) warrants use of a large truncation window in the time domain, even if the function is narrow. Another

observation that can be made for real signals (the case of signals from speech, images, etc.) is that the spectrum $F(s)$ is Hermite [Castleman96][Vitterli92]. The left half of the spectrum is a mirror image of the right half of the spectrum. Therefore, from the point of view of information content, the $N/2$ real and the $N/2$ imaginary values in the left half of the spectrum are redundant.

3.3.2 Filtering and Convolution

The most common operation in signal processing is filtering. Digital filtering can be implemented in at least two ways.

1. The filtering operation could be implemented by digital convolution of the sampled signal $x[n]$ and the filter coefficients $h[n]$.
2. The second method involves transforming of the input signal $x[n]$ and the filter impulse response $h[n]$ to the frequency domain using the Fourier Transform and then performing a simple multiplication.

If either the signal sequence $x[n]$ or filter coefficient sequence $h[n]$ is of shorter duration, then digital convolution is computationally efficient. In wavelets, the filter coefficients are of finite length. They are called Finite Impulse Response filters. Digital convolution seems to be a better choice for digital filtering.

Most naturally occurring signals are band limited. If the input signal $f(t)$ is band limited to $s_m = \frac{1}{2\Delta t}$, the convoluted signal would also be band limited and with suitable interpolation, the signal can be reconstructed. However there is inherent truncation which destroys bandlimitedness and additionally causes some aliasing. The effects of aliasing can be reduced to tolerable levels by over sampling, or by low pass filtering prior to sampling.

From digital signal processing techniques, we know that convolution with the impulse response produces artifacts at both ends of the truncation window by smearing the discontinuity. While the smearing effect at the ends of the truncation window cannot be avoided completely, it can be reduced to tolerable levels by :

1. making the truncation window wide with respect to the important components of the signal, so that nothing of interest is damaged.
2. arranging for $x[n]$ to have equal amplitude at each end of the truncation window, so that little or no discontinuity appears when it becomes periodic. This is usually done by multiplying the truncated function by a windowing function [Castleman96][Vitterli92] .

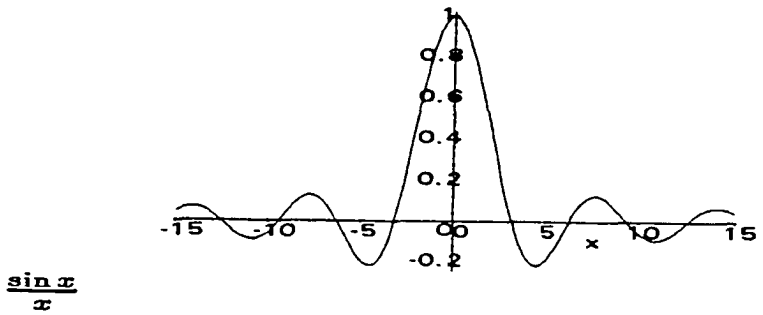


Figure 3-5 : Shows a sinc function.

The box function shown in Figure 3-6 is also the representation of an ideal low pass filter. Sub-optimal filters particularly those that are easy to implement by computer (the Haar or box function filter) can introduce artifacts into an image, usually without warning. Filters involving the rectangular pulse in one domain, have an unsavory behavior in the other

domain due to the infinite waves of the sinc $\left(\frac{\sin x}{x}\right)$ function producing a "ringing" effect. The sinc function and its Fourier are shown in Figures 3-5 and 3-6.

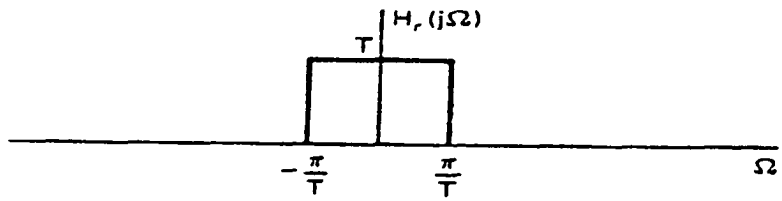


Figure 3-6 : Spectrum of the sinc function - ideal low pass filter

3.3.3 Down Sampling

Signal processing sometimes requires down sampling and up sampling of the signal. Down sampling [Aravena95][Oppenheim89][Strang96] is the process where alternate sample points are discarded and a new signal $v[n]$ is created. The index for the new sequence is the same as before but the sequence values correspond to every other point of the previous signal. This can be represented by the transformation in the spatial domain as shown below:

$v[n] = (\downarrow 2)x$ or $v[n] = x[2n]$ Eq. 3-29

In the Frequency Domain down sampling causes the Frequency response magnitude to be scaled down by a factor of 2. This is exactly the amount by which the sampling rate is reduced and the Frequency Spectrum gets "spread-out" by a factor of 2. Frequencies are doubled when compared to the original sequence.

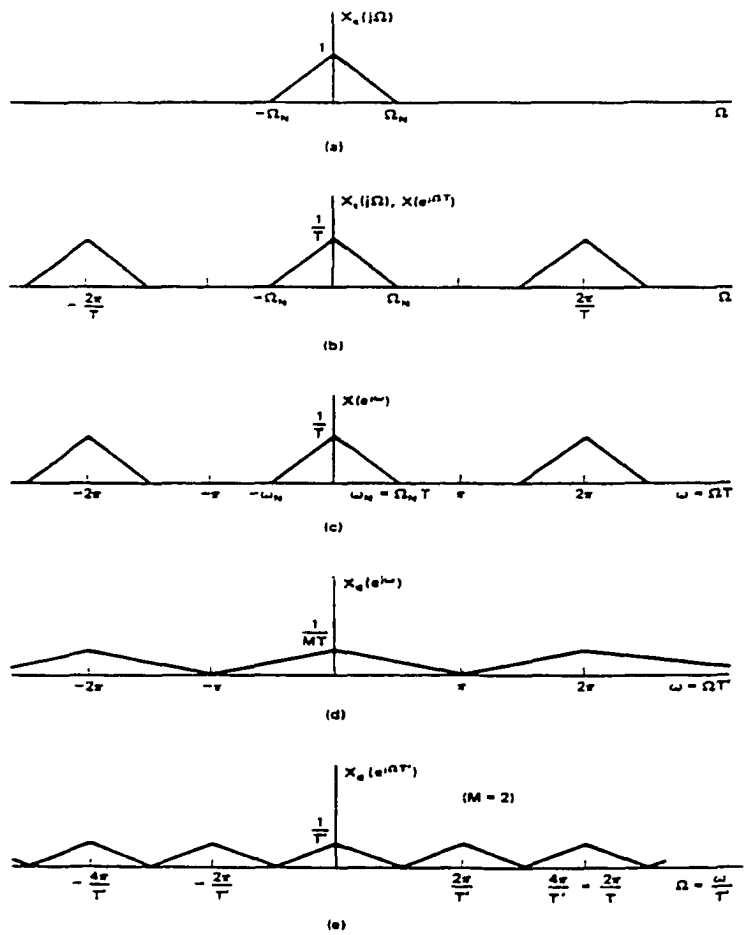


Figure 3-7 : Downsampling

(a) Figure 3-7(a) shows the Fourier Transform of a bandlimited continuous-time signal;
(b) Figure 3-7(b) shows the Fourier Transform of the impulse train of samples obtained with sampling period T;
(c) Represents the Fourier Transform of the impulse train of samples just as in Figure 3-7(b) but here the we are changing the frequency axis from Ω to ω;
(d) Figure 3-7(d) shows the discrete-time Fourier Transform of the downsampled by 2 with frequency represented as ω ;
(e) Figure 3-7(e) shows the same downsampled signal as in Figure 3-7(d) but here the frequency axis is in continuous time by Ω.

Mathematically this translates to:

$$V(w) = \frac{1}{2}\left[F\left(\frac{w}{2}\right) + F\left(\frac{w}{2} + \pi\right)\right] \qquad \text{Eq. 3-30}$$

Figure 3-7 [Oppenheim89] shows the Fourier Transform of a bandlimited continuous-time signal and the Fourier Transform of the sampled signal with sampling period T. Down sampling is illustrated in the Figure 3-7(d) and 3-7(e). In image processing this means that we examine the signal with a lower resolution. Down sampling can also cause aliasing if the highest frequency in the band limited signal w_n does not satisfy the relation $w_n M < \pi$, where M is the down sampling rate (in the previous example M was 2). Down sampling reduces the signal length in the time domain by half if the down sampling rate is 2. Down sampling is a common signal processing technique. In image processing applications it allows for lowering the image resolution. Since the filters in the Wavelet Analysis typically remains the same, application of the filter in a iterative fashion leads to frequency banding.

3.3.4 Up sampling

Up sampling [Aravena95][Oppenheim89][Strang96] is the process when the sampling rate is increased by a factor L. For purposes of simplicity, let us consider $L = 2$. Typically a zero is inserted between every other sample thus increasing the number of sample points by 2. Figure 3-8 [Oppenheim89] illustrates the up sampling process in the frequency domain. This is a process that (in some ways) is considered the opposite of down sampling. Mathematically this is represented as:

$$v[n] = \left[\uparrow 2\right] x[n] \qquad \text{Eq. 3-31}$$

$$v[2n] = x[n] \qquad \text{Eq. 3-32}$$

$v[2n+1] = 0$ Eq. 3-33

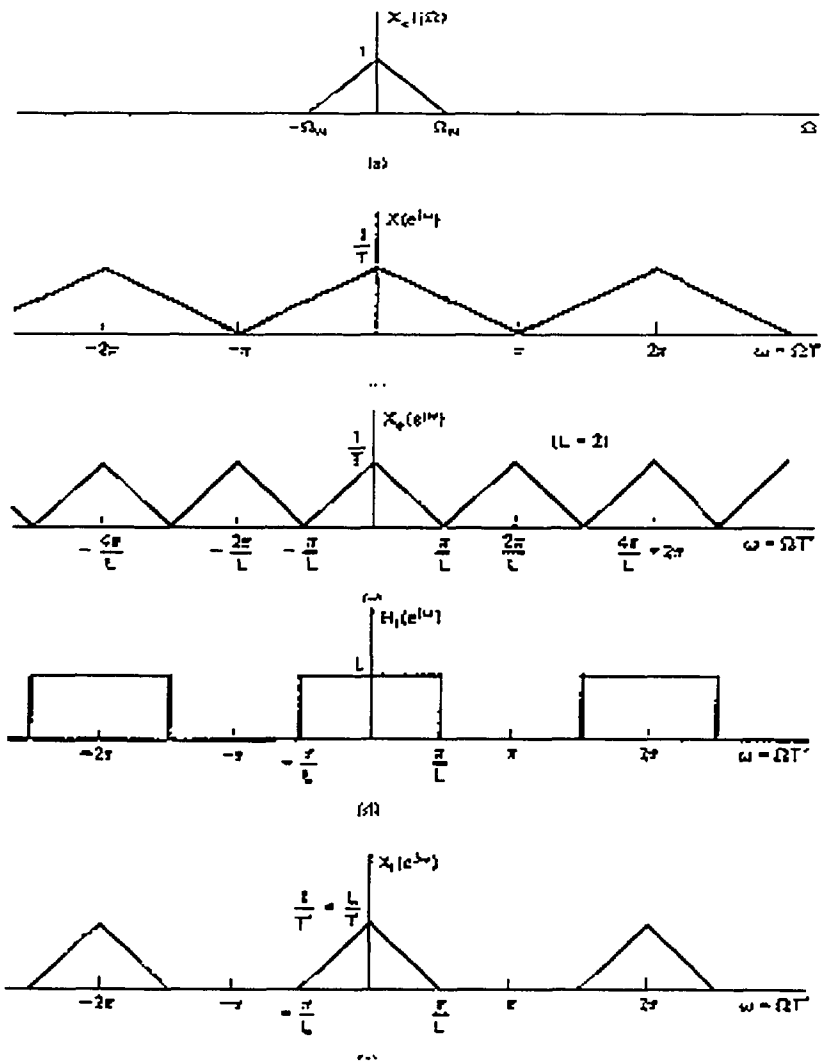


Figure 3-8 : Upsampling

a) Figure 3-8(a) shows the continuous-time Fourier Transform of a bandlimited signal.
b) Figure 3-8(b) shows the discrete-time Fourier Transform of the sampled bandlimited signal.
c) Figure 3-8(c) shows the Fourier Transform of the signal in Figure 3-8(b) down sampled by 2 .
d) Figure 3-8(d) shows the Fourier Transform of a ideal low pass filter with gain of 2 and cutoff frequency of $\frac{\pi}{2}$.
e) Figure 3-8(e) shows the Fourier Transform of the original signal after it subjected to upsampling and then filtered with the low pass filer shown in Figure 3-8(d).

In the Frequency domain this translates to the magnitude of the Frequency response of the original signal being scaled up by a factor of L, which in this case is 2. The Frequency spread shrinks by a factor of L and the frequency w_n becomes $\frac{w_n}{L}$. There are L copies of the spectrum in the range $[0,2\pi]$. This can be represented as:

$$V(w) = F(2w) \qquad \text{Eq. 3-34}$$

Instead of inserting zeros between every other sample point, usually a linear interpolation is done. If the intermediate points are fit with a Gaussian curve, then the Gaussian curve is used instead of the zeros in the upsampling process. This filtering leads to the "Gaussian smoothing". "Gaussian Smoothing" is commonly done in Medical Imaging where a three dimensional image is generated from two dimensional MRI or CT slices. The inter-slice points are obtained by "Gaussian smoothing". Most of the commercial signal processing software provide an upsampling function. The software must be used with a caution because they may actually be an interpolating function, instead of upsampling with zero.

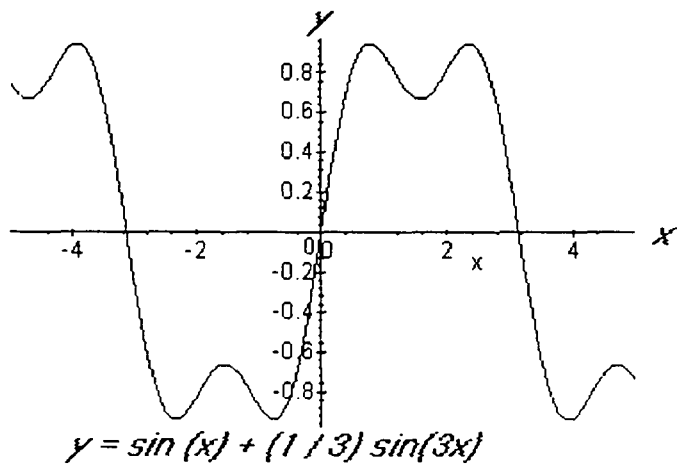


Figure 3-9 : Figure shows a plot of a simple sine wave

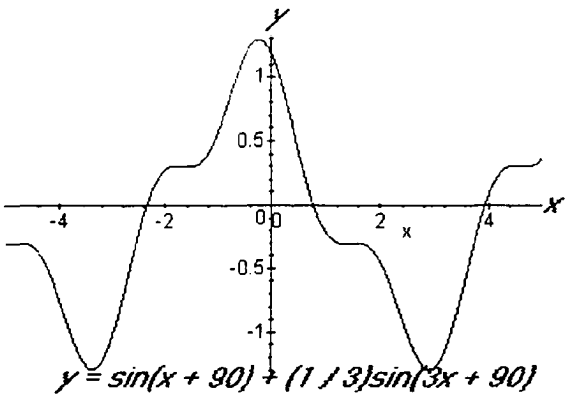


Figure 3-10 : The above figure shows the sine wave in Figure 3-9 with a non-linear phase added.

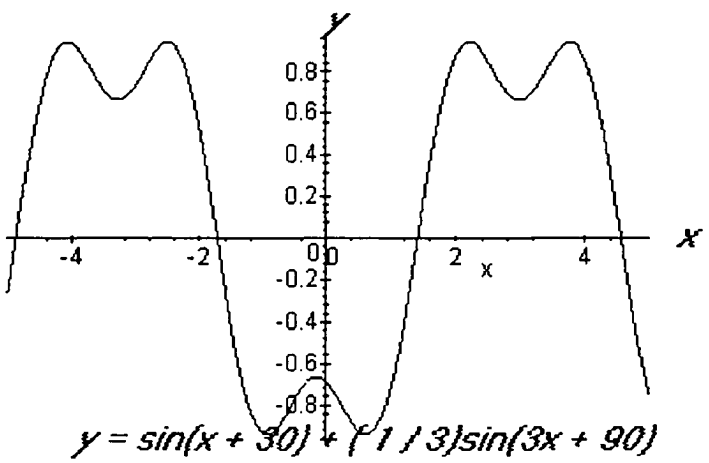


Figure 3-11 : The above figure shows the sine wave in Figure 3-9 with a linear phase added.

3.4 Frequency Domain Analysis

The Fourier Transform yields a break down of the frequency components of the signal. Another way to say this is, that the Fourier Transform yields the frequency magnitude response and a phase response in terms of the frequency. These characteristics are a function of only the frequency. They provide no idea about how the frequency changes

with respect to time. The Frequency Magnitude Response contains information about the intensity and the phase contains information about the waveform. Linear phase systems introduce what is called an ideal delay but preserve the wave form. Figure 3-9 shows a simple wave. Figures 3-10 and 3-11 show the same wave as in Figure 3-9 with a non-linear and linear phase added to it, respectively. From Figures 3-9, 3-10 and 3-11 we can observe that linearity preserves waveform. Linearity of phase causes the most acceptable form of signal distortion which makes signal reconstruction easier. It is for this reason that digital filters are usually designed with a linear phase.

3.4.1 Fourier Transform

The Fourier Transform [Vitterli92][Weaver83][Young93] is very useful for stationary signals. The Fourier Transform is a technique for representing a signal in terms of its spectrum. In signals that involve music, speech, images, etc., the spectrum evolves over time in a significant way. Their frequencies are constantly changing. This time evolution of these frequencies is not reflected in the Fourier Transform, at least not directly. The Fourier Transform contains information about the frequencies of the signal over all times (duration of the signal) instead of showing how the frequencies vary with time. Thus, there is no localized information when the Fourier Transform is considered.

For signals such as images with sharp variations, or signals with sharp discontinuities, a problem with the Fourier Transform is that in order to accommodate a discontinuity or a "steep-climb", high frequency terms appear that are not localized but are added everywhere.

This limitation led to the introduction of two-dimensional signal processing tools which provided [Kaiser94][Chan93]:

- time-frequency representations. For example, the narrow Band Ambiguity Function (AF), Wigner Distribution, Short Time Fourier Transform, etc., provided the spectral properties that are time dependent, i.e., non-stationary or quasi-stationary.
- time-scale representations. For example, the wide Band Ambiguity Function, Wavelet Transform, etc., extract the localized contributions of the signals with the scale parameter.

The delta function $\delta(t)$, has an infinite temporal resolution and no spectral resolution as illustrated in Figure 3-12. Figure 3-12 also shows the Fourier Transform or spectrum of a time limited function.

The Fourier Transform is slave to the Sine and Cosine functions. When we compute the Fourier Transform, we are actually representing the function in terms of Sine and Cosine functions which are periodic functions. The Fourier Transform is actually a great tool for analyzing stationary signals. However they are inadequate when it comes to representing non-stationary signals. Gabor Transforms attempted to address this problem. The Fast Fourier Transform is simply a more computationally efficient reformulation of the Fourier Transform.

3.4.2 Walsh Transform

The difficulty of the frequency spectrum being unable to provide a time component can be overcome with the Walsh transform. This transform has properties similar to the Fourier Transform, but with piece-wise constant bases instead of the exponential sine and cosine bases [Fournier94].

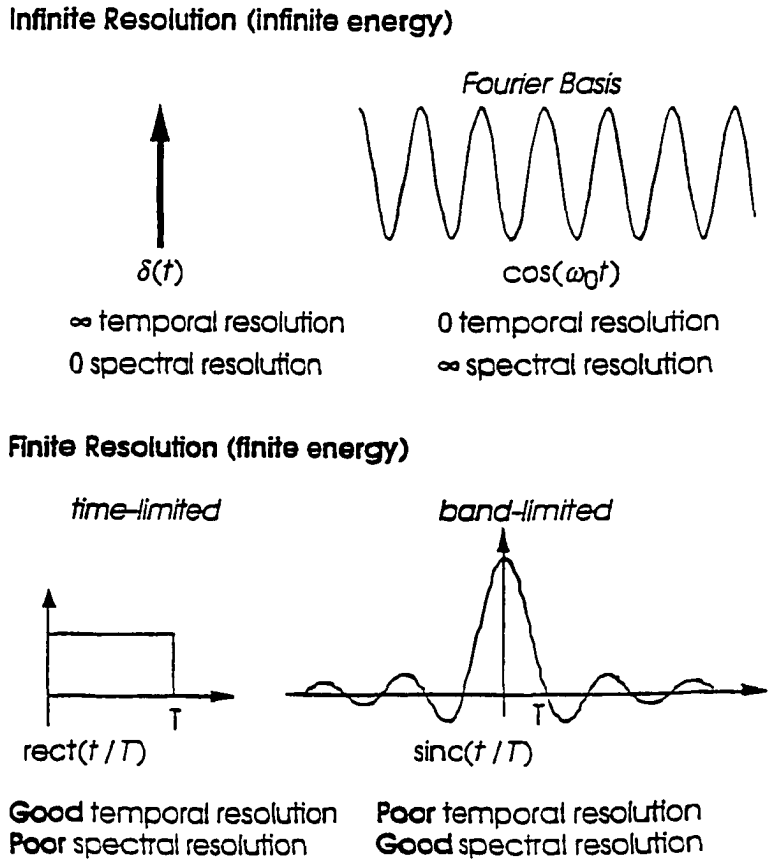


Figure 3-12 : Shows figures of functions with poor and good spectral resolution[Orr96].

The Walsh Transform coefficients are given by:

$$w_i = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(t) W_i(t) dt \qquad \text{Eq. 3-35}$$

where

$$W_{2i-q}(t) = (-1)^{\frac{i}{2}-q} \times \left[W_i(2t+\frac{1}{2}) + (-1)^{(i-q)} W_i(2t-\frac{1}{2}) \right] \qquad \text{Eq. 3-36}$$

with $w_0(t) = 1$, where j ranges from 0 to ∞, and $q = 0$ or 1.

The Walsh functions are normally defined $-\frac{1}{2} \le t \le \frac{1}{2}$ and are 0 outside of this interval. The orderings for their index i is usually equal to the number of times the function changes sign (zero-crossings of the function). Figure 3-13 is a time domain representation of the first 8 Walsh functions.

The Walsh function can successfully be used to represent certain discontinuities if they occur at integral values of the spatial or time variable. But for the vast majority of signals, where discontinuities may occur anywhere in the signal, the Walsh Transform introduces high frequency terms that are added globally for all t to deal with singularities. Therefore, the Walsh Transform suffers from the same deficiency as other transforms of unlocalized information.

3.4.3 Short Time Fourier Transform

The STFT [Chui][Rioul91][Kadambe92] attempts to deal with the non-stationary aspect of signals by chopping the signal or windowing the signals into short consecutive segments and then computing the Fourier Transform of the Windowed Signal. In the Gabor Transform a Gaussian Window is used.

The STFT of the signal $f(t)$ is defined as:

$$STFT(w,\tau) = \int_{-\infty}^{\infty} f(t) h^*(t-\tau) e^{-iwt} dt \qquad \text{Eq. 3-37}$$

where $h(t)$ is the sliding window function centered at time τ.

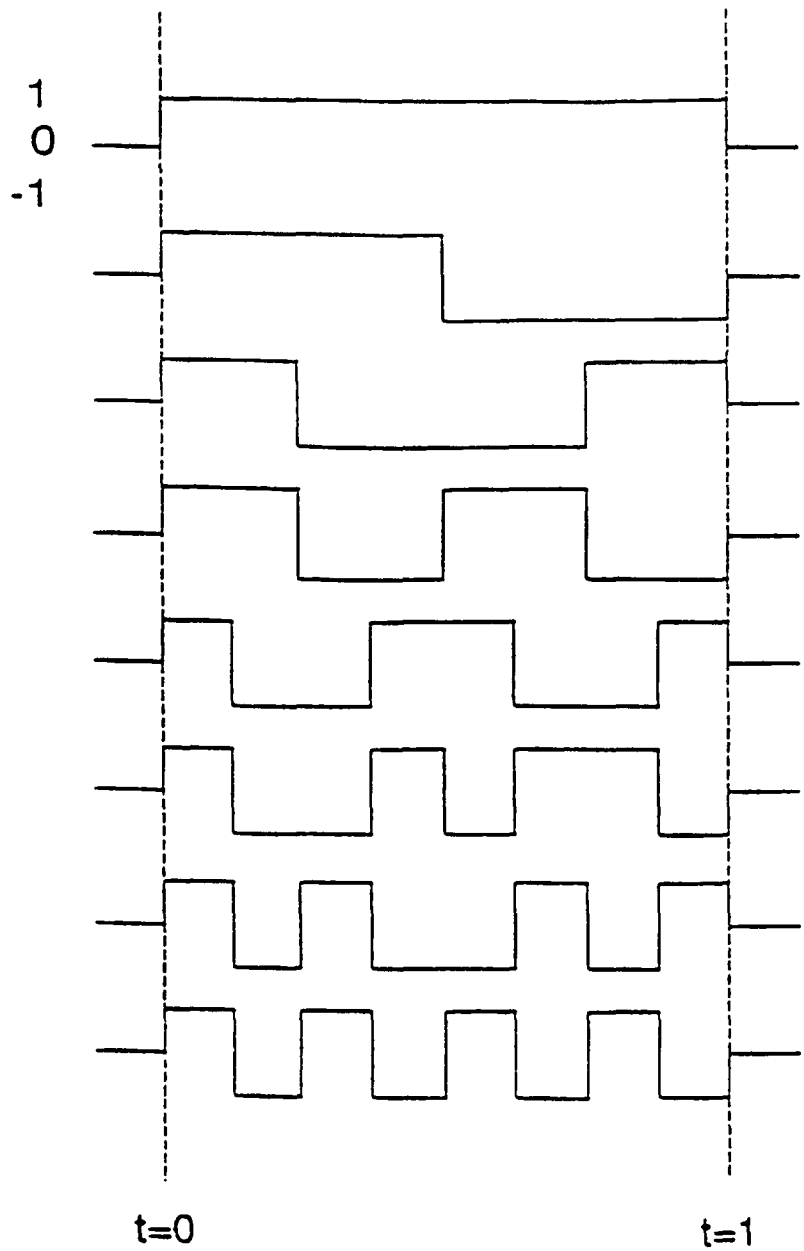


Figure 3-13 : Figure shows Walsh functions [Fournier94]

In mathematical terms this equation is an expansion of the signal in terms of a family of functions $h^*(t-\tau)e^{iwt}$, which are generated through translations τ in time and translations w in frequency. In the case of the Gabor Transform [Gabor46] the sliding window is a Gaussian Window of the form

$$h(t) = h(0)e^{-at^2} \text{ and } STFT(w,\tau) \in L^2(R^2) \qquad \text{Eq. 3-38}$$

This transform is often used in signal processing especially in the processing of speech signals. This implementation is done by convolving the original signal with a constant shape window function like a Gaussian window, computing the Fourier Transform, and then repeating the process traversing across the length of the signal. The results can be plotted in a time-frequency plane where the size of the window remains the same throughout the plane. One known problem with this method is that it does not capture frequency changes that occur inside the window. One way to capture frequency changes with time is to decrease the size of the windowing function. However, this results in too many windows and increased computational time. Additionally, the lower frequency components are poorly resolved. It is not possible to obtain arbitrarily fine localization in time and in frequency due to the uncertainty principle. The uncertainty principle dictates the lower bound for the time-bandwidth product of possible basis functions by the equation

$$\Delta t \bullet \Delta w \geq \frac{1}{4\pi} \qquad \text{Eq. 3-39}$$

where (Δt) and (Δw) are variances of the absolute values of the function and its Fourier Transform respectively [Chui][Araven95].

The Fourier Transform provides information regarding how the amplitude changes with frequency or the phase changes with frequency. The windowed Fourier Transform (STFT)

decomposes a function into a set of frequency intervals of the same size in the frequency domain. The spatial and frequency resolution of a Windowed Fourier Transform is constant. In most naturally occurring signals like speech, images, seismic signals, music, etc., high frequency components are of shorter duration and low frequency components are of longer duration. With a single constant window, it is difficult to analyze both the fine and large structures of an image. The single window resolution also introduces misleading high frequencies when decomposing local features Mallat89][Mallat92.1] [Mallat92.2].

Consider a function $f(t)$ that changes from a value of 0 to 1 in a very short interval of time Δt in the neighborhood of some point t_0. The function has a value of 0 prior to the edge t_0 and has a value of 1 after the edge. Additionally assume that the rise of the edge, in the neighborhood t_0, is a pure sinusoid wave of frequency w_0. It is reasonable to expect that the STFT coefficients would decrease rapidly for terms with frequencies greater than w_0. However, in reality, this does not happen because the edge is very localized and has only half of the sinusoidal wave's period. Consequently when the frequency w is large with respect to w_0, the modulus of the STFT coefficients $S(t_0, w)$ decreases slowly. Although the signal $f(t)$ is locally a pure sinusoidal wave of frequency w_0, at a frequency $2w_0$, the STFT coefficient $|S(t_0, w_0)|$ is still about half the value of $|S(t_0, 2w_0)|$ according to Mallat. This numerical property makes it hard to interpret the Windowed Fourier coefficients when the feature is very localized with respect to the size of the support window function $h(t)$. From a computational standpoint, the inability of the STFT to produce an orthonormal decomposition also makes it less desirable [Chan93] [Young93] [Rioul92].

3.4.4 Wigner-Ville Distribution

Wigner-Ville distribution was first developed for characterization of phase-space uncertainty in Quantum Mechanics [Wigner32]. Since phase-space uncertainty and time-frequency uncertainty are analogous, it is a tool for time-frequency analysis [Classen80].

The Wigner-Ville distribution is a time-varying power spectral density function computed as a time-varying, instantaneous estimate of autocorrelation [Kadambe92]. The Wigner-Ville distribution attempts to overcome the time versus frequency trade off of the STFT with a distribution that can be represented as

$$W(t,w) = \int_{-\infty}^{\infty} f\left(t+\frac{\tau}{2}\right) f^*\left(t-\frac{\tau}{2}\right) e^{-2i\pi w\tau} d\tau \qquad \text{Eq. 3-40}$$

where f^* is the complex conjugate of f. This corresponds to the Weyl-Heisenberg coherent state arising from translations and modulations of a single function.

This distribution exhibits a number of useful properties like time-frequency support or nonpositivity, instantaneous frequency or beat frequencies (being real valued and perfectly localized for chirp signals)[Kaiser94], group delay, etc. By its bilinear nature, it introduces cross-correlation terms in the time-frequency energy representation. This makes interpretations of the coefficients difficult. Its Wavelet Transform also exhibits cross terms in the energy distribution, however this is overcome by choosing the wavelet such that the Fourier Transform of the wavelet $F(w) = 0$ for $w < 0$.

3.5 Wavelets and Filter Banks

As quoted in the literature [Viterli92][Mallat97][Duchowski98] that a constant logarithmic scale for frequency is natural to the ear and natural to the human eye. So the preference would be to observe high frequency components in a highly localized time window and to observe the low frequency components with longer time windows or windows with better

frequency resolution. The Wavelet Transform being a function of scale and space provides this opportunity. Performing affine operations on the single prototype wavelet, usually referred to as "mother wavelet" in wavelet literature, creates a set of scaled and translated versions of the mother wavelet called the wavelet basis set and this set serves as a set of basis functions for the analysis and synthesis of functions. Mathematically, the Continuous Wavelet Transform (decomposition) of a function $f(t)$ with respect to a "mother wavelet" $\psi_{a,b} = {}^{1}\!/\!_{\sqrt{a}}\,\psi\left(\frac{t-b}{a}\right)$ for $a>0$ can be written as inner products in $L^2(R)$ as shown :

$$WT[f](a,b) = < f(t), \psi\left(\frac{t-b}{a}\right) > = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \qquad \text{Eq. 3-41}$$

where $a,b \in R$.

- The factor ${}^{1}\!/\!_{\sqrt{a}}$ ensures that the functions $\psi_{a,b}$ have a constant norm in the space $L^2(R)$ of square integrable functions.

Here the original function $f(t)$, a function over the independent variable t, is mapped into a new two-dimensional function across scale a and translation b. A wavelet coefficient $WT[f](a,b)$ at a particular scale and translation represents how well the signal, $f(t)$, and the scaled and translated $\psi(t)$ "mother wavelet" match. Therefore, if a part of the signal $f(t)$ is similar to the "mother wavelet" at a certain scaling and translation, then the wavelet coefficient at that scale and translation, $WT[f](a,b)$, will have higher magnitude. In other words, the coefficient represents the amount of correlation between the signal and the wavelet at a particular scale and translation.

Since wavelets provide a finite impulse response filter, convolution in the time domain is a way to obtain the wavelet coefficients. The bandwidth of the filters is proportional to $\frac{1}{a}$ [Prasad97][Strang96][Young93]. The scale parameter a is normally chosen such that it is inversely proportional to the frequency w to obtain a time-frequency representation [Kadambe92][aravena96][Strang96][Young93]. Since the scale a is chosen to be inversely proportional to the frequency w, the Wavelet Transform provides high spectral resolution and poor temporal resolution for low frequency. By the same argument, when a is small the Wavelet Transform provides high temporal resolution enabling a zooming in effect of the singularities, and other local features of the signal.

The remainder of this section addresses the types of functions that are admissible as wavelets and their relation to multiresolution analysis.

Wavelets are localized waves or functions with mean zero. For a function to be a wavelet it must satisfy the following conditions [Aravena95][Chan93] [Kaiser94][Young93]:

1. the wavelet must have compact support.
2. $\int_{-\infty}^{\infty} \psi(t)dt = \hat{\psi}(0) = 0$. This condition implies that the zero frequency component must be zero. $\hat{\psi}(w)$ is the Fourier Transform of $\psi(t)$. This condition implies that the wavelet exhibits an oscillatory behavior, i.e., the function takes on both positive and negative values.

Figure 3-14 provides an example of a wavelet namely the Haar Wavelet and its spectrum. Figure 3-14 also includes the scaling function and its Fourier Spectrum for a Daubechies wavelet with 8 vanishing moments.

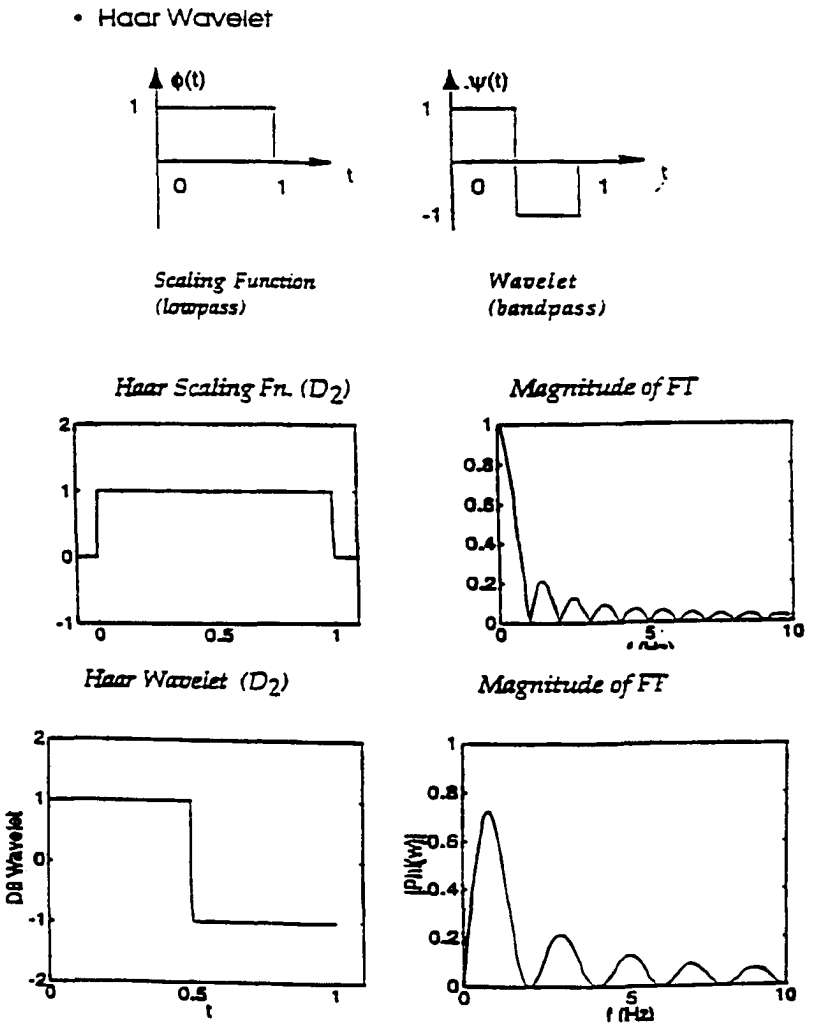


Figure 3-14 : This figure shows the Haar scaling function and Haar wavelet along with the Fourier spectrum of each of these functions. The figure also shows the Duabechies D8 wavelet and its Fourier spectrum [Orr96].

For computational efficiency integral powers of 2 are used for frequency partitioning and this provides a binary dilation of 2^j, a dyadic translation of $\frac{b}{2^j}$, and a typical wavelet set

obtained from $\psi(2^j x - k) \forall j,k \in Z$. This is normalized to a unit magnitude and yields basis functions of the form

$$\psi_{j,k}(t) = 2^{\frac{j}{2}} \psi\left(2^j x - k\right). \qquad \text{Eq. 3-42}$$

This basis will constitute an orthonormal basis if

$$\left\langle \psi_{j,k}(t), \psi_{m,n}(t) \right\rangle = \delta_{jk} \bullet \delta_{mn} \qquad \text{Eq. 3-43}$$

δ is the Kronecker delta.

Any $L^2(R)$ function can be approximated, up to as arbitrarily small precision by a finite linear combination of the $\psi_{a,b}$ [Daubechies88][Daubechies93].

The Wavelet transform $WT[f](1/2^j, k/2^j)$ yields an integral wavelet transform evaluated at dyadic position $b = k/2^j$ with binary dilation $a = 2^{-j}$ [Orr96]. The existence of the inverse wavelet transform depends on the choice of the mother wavelet ψ. Mathematically for the inverse transform to exists, the wavelet must satisfy an admissibility condition

$$0 < C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(w)|^2}{|w|} dw < +\infty. \qquad \text{Eq. 3-44}$$

The reconstruction formula for the function f is given by

$$f(t) = \frac{C_\psi^{-1}}{\sqrt{a}} \int_0^{\infty} \frac{da}{a^2} \int_{-\infty}^{\infty} WT[f](a,b) \psi\left(\frac{t-b}{a}\right) db \qquad \text{Eq. 3-45}$$

for $a > 0$. It is these conditions on C_ψ that require $\hat{\psi}(0) = 0$ and that ψ oscillates. If ψ oscillates then $\int_{-\infty}^{\infty} \psi(t) dt = 0 \Rightarrow \hat{\psi}(0) = 0$. The Continuous Wavelet Transform is more like a group of Short Time Fourier Transforms with a different window for each

frequency. The STFT, (refer to Eq. 3-38) are obtained by translations through time and translations in frequency thus providing a constant bandwidth and generating a spectrogram with fixed spectral and temporal resolution. In contrast, the Wavelet Transform is an expression in functions of $\frac{1}{\sqrt{a}}\psi\left(\frac{(t-b)}{a}\right)$ generated by translations of b in time and dilations of a in scale.

Both the STFT and the Wavelet Transform are highly redundant when the parameters (w,τ) and (a,b) are continuous. Therefore, the transforms are evaluated on a discrete grid on the time-frequency and time-scale plane. Another reason for this is to reduce the required computational time. When this grid allows a set of basis functions to be an orthonormal basis, we get a parsimonious representation meaning to say there is no redundancy. This is a very natural process for signals obtained from images. Usually we consider a dyadic grid; however, there is a good reason to consider a biorthonormal wavelet basis. Practical uses of a non-orthonormal basis is not common at this stage of evolution; however, permitting one and allowing redundancy for the purposes of error correction especially in the case of communications over network lines shows its importance.

When using Wavelet Transforms we will construct, in this research, wavelets or wavelet coefficients that lead to perfect reconstruction, especially when quantization errors are not considered. These filters in the analysis and synthesis banks are anti-aliasing meaning they cancel aliasing. To make matters easy, quadrature mirror filters may be added as a condition to make the construction of the synthesis bank filters from the analysis filters (or vice versa).

From a signal processing point of view, a wavelet is a bandpass filter. In the dyadic case it is actually an octave band filter. The Wavelet Transform can be interpreted as a constant-Q filtering with a set of octave band filters, followed by filtering at the respective Nyquist frequencies (corresponding to the bandwidth of that particular frequency). Based on this interpretation the multiresolution analysis view was introduced by Mallat [Mallat89] [Mallat92.1] and Meyer [Meyer92]. This multiresolution view is essentially a successive approximation procedure where we consider the signal at different levels of resolution. This process is widely used in image compression, computer vision and signal processing. The wavelet decomposition is a successive approximation method which adds more and more projections onto detail spaces or spaces spanned by wavelets and their shifts at different scales.

3.6 Conclusions

Wavelets can be used as a tool for detecting singularities. In this dissertation, wavelets will be used to detect the singularities. The location of the singularities leads to a set of points that are then employed, to register medical images. How this is done forms the contents of the next chapter. The next chapter dwells into the two dimensional algorithm to find the wavelet coefficients where in one case the image is down sampled at each stage of the iteration and in the other case the filter coefficients are upsampled instead of the down sampling of the image matrix.

CHAPTER 4

WAVELET MODULUS MAXIMA IMAGE

4.1 Introduction

Singularities and irregular structures carry important information about signals. Images with sharp changing intensity or gray scale provide the locations of the object contours. This is useful in the recognition of edges, boundaries and important features. Decomposing signals with the Wavelet Transform into elementary building blocks that are localized both in space and frequency, can characterize the local regularity of signals. The local regularity of a function is normally measured with Lipschitz exponents and these can be used for signal characterization.

4.2 Definitions and Derivations

The information presented in this section is purely for information and background purposes, so that the reader would have a better comprehension of the research. The information that forms the initial part of this section comes from various sources and is mostly what would be the content of a senior level or graduate course on "Real Analysis" [Aravene96][Burden93][Mallat92.2][Cohen89][Royden88]. However the content is presented in the context of its application to this research. We have also presented a more detailed derivation of some of the results because the current reviewed literature assumes these results to be trivially obvious. At the early stages of this research, it was felt that the Lipshitz exponents could be also be used in our technique to register images. However as we developed our registration method, we realized that it was not significant to our method. Therefore the technique we developed does not use Lipschitz exponents. The

registration technique developed in this research uses the Wavelet Modulus Maxima technique developed by Mallat [Mallat89] [Mallat92] and Zhong [Zhong90]. The theoretical issues relating to the Wavelet Modulus Maxima is discussed in section 4.3.

- A function $f:[a,b] \to R$ is said to satisfy a Lipschitz condition with Lipschitz constant L on $[a,b]$ if for every $x, y \in [a,b]$,

  $$|f(x) - f(y)| \le L|x - y| \qquad \text{Eq. 4-1}$$

- If f satisfies a Lipschitz condition with Lipschitz constant L on an interval $[a,b]$ then

  $$f \in C[a,b] \qquad \text{Eq. 4-2}$$

  i.e., f is continuous in the closed interval $[a,b]$.

- Let f be a function defined on a set X of real numbers and $x_0 \in X$. f is said to be continuous at x_0 if

  $$\lim_{x \to x_0} f(x) = f(x_0) \qquad \text{Eq. 4-3}$$

  The function f is said to be continuous on X if it is continuous at each number in X. $C(x)$ denotes the set of all continuous functions on X.

- If f has a derivative that is bounded on $[a,b]$ by L, then f satisfies a Lipschitz condition with Lipschitz constant L on $[a,b]$.

- A function $f(t,y)$ is said to satisfy a Lipschitz condition in one variable y on a set $D \in R^2$, provided a constant $L > 0$ exists with the property

  $$|f(t,y_1) - f(t,y_2)| \le L|y_1 - y_2| \qquad \text{Eq. 4-4}$$

whenever $(t,y_1),(t,y_2)\in D$. L is called the Lipschitz constant for f. If for $L>0$

$$\left|\frac{df}{dy}(t,y)\right|\le L \qquad \text{Eq. 4-5}$$

for all $(t,y)\in D$ then f satisfies a Lipschitz condition on D in the variable y with Lipschitz constant L.

From elementary calculus we know that the continuity of a function at a point does not imply differentiability of the function at that point. However, differentiability does imply continuity.

- Let n be a positive integer and let α be such that $n\le\alpha\le n+1$. A function $f(x)$ is said to be Lipschitz α at x_0, if and only if there exists two constants $A>0$ and $h_0>0$ and a polynomial of order n, $P_n(h)$, such that for $h<h_0$

$$\left|f(x_0+h)-P_n(h)\right|\le A|h|^{\alpha} \qquad \text{Eq. 4-6}$$

- The function $f(x)$ is uniformly Lipschitz α over the interval (x_1,x_2), if and only if there exists a constant $A>0$ and for any $x_0,x_0+h\in(x_1,x_2)$ there exists a polynomial of order n, $P_n(h)$ such that

$$\left|f(x_0+h)-P_n(h)\right|\le A|h|^{\alpha} \quad . \qquad \text{Eq. 4-7}$$

- Lipschitz regularity of $f(x)$ at x_0, is the superior bound of all values α such that $f(x)$ is Lipschitz α at x_0.
- A function is singular at x_0, if it is not Lipschitz 1 at x_0.
- If $f(x)$ is Lipschitz α, for $\alpha>n$, then $f(x)$ is n times differentiable at x_0 and the polynomial $P_n(h)$ is the first $n+1$ terms of the Taylor series of $f(x)$ at x_0.

Note that for $n = 0$, $P_n(h) = f(x_0)$.

Suppose $f \in C^n[a,b]$, $f^{(n+1)}$ exists on $[a,b]$ and $x_0 \in [a,b]$. For every $x \in [a,b]$ there exists $\xi(x)$ between x_0 and x with

$$f(x) = P_n(x) + R_n(x) \qquad \text{Eq. 4-8}$$

where

$$P_n(x) = f(x_0) + f^{(1)}(x_0)(x - x_0) + \frac{f^{(2)}(x_0)(x - x_0)^2}{2!} + \ldots + \frac{f^{(n)}(x_0)(x - x_0)^n}{n!} \qquad \text{Eq. 4-9}$$

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)(x - x_0)^k}{k!} \qquad \text{Eq. 4-10}$$

and

$$R_n(x) = \frac{f^{(n+1)}(\xi)(x - x_0)^{n+1}}{(n+1)!} \qquad \text{Eq. 4-11}$$

The Taylor polynomial [Burden93] is also called a Maclaurin Polynomial (if $x_0 = 0$) and that Taylor series is called the Maclaurin series.

- The Lipschitz regularity α_0 gives an indication of the differentiability of $f(x)$. If the Lipschitz regularity α_0 of $f(x)$ where $n < \alpha_0 < n+1$ is n times differentiable at x_0 but its n^{th} derivative is singular at x_0 then α_0 characterizes this singularity.
- If $f(x)$ is Lipschitz then its primitive $g(x)$ is Lipschitz $\alpha + 1$. However it is not true that if a function is Lipschitz α at a point x_0, then its derivative is Lipschitz $\alpha - 1$ at the same point.
- If α is not an integer and $\alpha > 1$, a function is uniformly Lipschitz α on an open interval (x_1, x_2), if and only if its derivative is uniformly Lipschitz $\alpha - 1$ on the same interval.

The local maxima of the Wavelet Transform modulus provides enough information for analyzing singularities [Lu94][Lu96][Holshneider93][Mallat92]. These authors prove that the modulus maxima detects all singularities and the local Lipschitz exponents can be measured from their evolution across scales. The local frequency of the oscillations of a digital signal can be measured from the points where the modulus of the Wavelet Transform is a local maximum with respect to both the scale and spatial variable [Rioul92][Zhong90]. By analyzing the Wavelet Transform Modulus Maxima across scales, the white noise present in the signal can be differentiated from the original image information.

The classical tool for measuring the Lipschitz regularity of a function $f(x)$ is to evaluate the asymptotic decay of its Fourier Transform $F(w)$. A bounded function $f(x)$ is uniformly Lipschitz α over R if it satisfies

$$\int_{-\infty}^{\infty} |F(w)|\left(1+|w|^{\alpha}\right)dw < +\infty \qquad \text{Eq. 4-12}$$

where

$$F(w) = \int_{-\infty}^{\infty} f(x)e^{-iwx}dx \qquad \text{Eq. 4-13}$$

This condition is sufficient but not necessary. This condition provides a global regularity condition over the whole real axis, however from this condition it is not conclusive whether the function is locally more regular at any particular point x_0. This is because the Fourier transform unlocalizes the information along its spatial variable. Hence, the Fourier transform does not measure the local Lipschitz regularity of a function efficiently or accurately.

If one assumes that a wavelet has compact support, then the value of the Wavelet Transform $WT[f](a, x_0)$ depends upon the values of $f(x)$ in a neighborhood of x_0. In addition, the magnitude is proportional to the scale a. At fine scales (i.e., for large values of a) the Wavelet Transform will yield localized information about $f(x)$. A Wavelet Transform does provide excellent localization in both time and frequency especially as contrasted to the Fourier Transform, the Gabor Transform or the Short Time Fourier Transform, Lapped Orthogonal Transforms, etc. [Holshneider] [Mallat92.][Strang96].

- The number of vanishing moments relates to the smoothness of a function. To discuss vanishing moments of functions, we first consider the vanishing moments of a wavelet which are nothing but functions which satisfy the certain conditions which has been discussed in chapter 3. A wavelet $\psi(x)$ is said to have n vanishing moments, if and only if for all positive integers $k < n$, it satisfies,

$$\int_{-\infty}^{\infty} x^k \psi(x) dx = 0. \qquad \text{Eq. 4-14}$$

As the number of vanishing moments increases in the time domain the oscillations of the function increase and in the frequency domain the function exhibits a smoother decay. The consequence of the cancellations of $\psi(x)$ is that the values of $WT[f](a, x_0)$ are influenced by the regularity of the function f. In the previous chapter the Wavelet Transform was defined as

$$WT[f](a,b) = \left\langle f(t), \psi\left(\frac{t-b}{a}\right) \right\rangle. \qquad \text{Eq. 4-15}$$

where if $C_\psi = \int \frac{|\psi(\omega)|^2}{\omega} d\omega < \infty$, its inverse transform is:

$$\cdot f = C_\nu^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T^{\omega}[f(x)](a,b)\psi\left(\frac{t-b}{a}\right) \qquad \text{Eq. 4-16}$$

This provides the basis of the following derivation, for a wavelet ψ with $n+1$ vanishing moments

$$\left|\left\langle f, \psi_{a,b}\right\rangle\right| = \left|\left\langle f - P + P, \psi_{a,b}\right\rangle\right| = \left|\left\langle f - P, \psi_{a,b}\right\rangle + \left\langle P, \psi_{a,b}\right\rangle\right| \qquad \text{Eq. 4-17}$$

$$\left|\left\langle f - P, \psi_{a,b}\right\rangle + \left\langle P, \psi_{a,b}\right\rangle\right| \le \left|\left\langle f - P, \psi_{a,b}\right\rangle\right| + \left|\left\langle P, \psi_{a,b}\right\rangle\right| \qquad \text{Eq. 4-18}$$

since $\left|\left\langle P, \psi_{a,b}\right\rangle\right| = 0$ whenever the degree of the inner product is $\left|\left\langle P, \psi_{a,b}\right\rangle\right| \le \alpha - 1$

$$\left\langle a_0 + ta_1 + t^2 a_2 + \ldots + t^n a_n, \psi_{a,b}\right\rangle = a_0\left\langle 1, \psi_{a,b}\right\rangle + a_1\left\langle t, \psi_{a,b}\right\rangle + a_2\left\langle t^2, \psi_{a,b}\right\rangle + \ldots + a_n\left\langle t^n, \psi_{a,b}\right\rangle$$

Eq. 4-19

and, $\left\langle 1, \psi_{a,b}\right\rangle = 0$, $\left\langle t, \psi_{a,b}\right\rangle = 0$, $\left\langle t^2, \psi_{a,b}\right\rangle = 0, \ldots \left\langle t^n, \psi_{a,b}\right\rangle = 0$. Eq. 4-20

Using these results and applying them to our equations we have

$$|WT[f](a,b)| = \left|\left\langle f, \psi_{a,b}\right\rangle\right| = \left|\left\langle f - P, \psi_{a,b}\right\rangle\right| \le C_1 \int_{-\infty}^{\infty} |t - t_0|^{\alpha} |\psi_{a,b}(t)| dt \qquad \text{Eq. 4-21}$$

$$C_1 \int_{-\infty}^{\infty} |t - t_0|^{\alpha} |\psi_{a,b}(t)| dt = C_1 \sqrt{a} \int_{-\infty}^{\infty} |ay + b - t_0|^{\alpha} |\psi(y)| dy \qquad \text{Eq. 4-22}$$

$$|ay + b - t_0|^{\alpha} \le C\left(|ay|^{\alpha} + |b - t_0|^{\alpha}\right) \le C\left(|a|^{\alpha}|y|^{\alpha} + |b - t_0|^{\alpha}\right) \qquad \text{Eq. 4-23}$$

$$C_1\sqrt{a} \int_{-\infty}^{\infty} |ay + b - t_0|^{\alpha} |\psi(y)| dy \le C_2\sqrt{a} \int_{-\infty}^{\infty} \left(|a|^{\alpha}|y|^{\alpha} + |b - t_0|^{\alpha}\right) |\psi(y)| dy \qquad \text{Eq. 4-24}$$

where $C_2 = CC_1$. Let $|b - t_0| \le C_3 a$. From this we have

$$|b - t_0|^{\alpha} \le |C_3|^{\alpha} |a|^{\alpha}. \qquad \text{Eq. 4-25}$$

We now have

$$|a|^{\alpha}|y|^{\alpha}+|b-t_0|^{\alpha}\le|a|^{\alpha}\left(|y|^{\alpha}+|C_3|^{\alpha}\right) \qquad \text{Eq. 4-26}$$

$$WT[f](a,b)|\le C_2\sqrt{a}\int_{-\infty}^{\infty}\left(|a|^{\alpha}|y|^{\alpha}+|b-t_0|^{\alpha}\right)|\psi(y)|dy \qquad \text{Eq. 4-27}$$

$$|WT[f](a,b)\le C_2\sqrt{a}\int_{-\infty}^{\infty}\left(|a|^{\alpha}|y|^{\alpha}+|C_3|^{\alpha}|a|^{\alpha}\right)|\psi(y)|dy \qquad \text{Eq. 4-28}$$

$$|WT[f](a,b)|\le C_2\sqrt{a}\int_{-\infty}^{\infty}\left(|a|^{\alpha}|y|^{\alpha}|\psi(y)|+|C_3|^{\alpha}|a|^{\alpha}|\psi(y)|\right)dy \qquad \text{Eq. 4-29}$$

$$|WT[f](a,b)|\le C_2\sqrt{a}\int_{-\infty}^{\infty}|a|^{\alpha}|y|^{\alpha}|\psi(y)|dy+C_2\sqrt{a}\int_{-\infty}^{\infty}|C_3|^{\alpha}|a|^{\alpha}|\psi(y)|dy \qquad \text{Eq. 4-30}$$

$$|WT[f](a,b)|\le|a|^{\alpha+(\frac{1}{2})}|K_1+K_2| \qquad \text{Eq. 4-31}$$

$$|WT[f](a,b)|\le K|a|^{\alpha+(\frac{1}{2})} \qquad \text{Eq. 4-32}$$

This shows that as the scale a goes to zero, the amplitude of wavelet coefficients $|WT[f](a,b)|$ in the region $|b-t_0|\le C_3a$ decays rapidly if f is regular at t_0, however if f has some singularity at this point, the decay is slower [Cohen96].

If $|F(w)|\le\dfrac{C}{(1+|w|)^{\alpha+\beta}}\Rightarrow\dfrac{d^{\alpha}}{dx^{\alpha}}f(x)$ exists for $\alpha<\beta$ and $\beta>1$

The Inverse Fourier Transform is given by the expression

$$f(x)=\int_{-\infty}^{\infty}F(w)e^{2\pi ixw}dw \qquad \text{Eq. 4-33}$$

Differentiating both sides we have

$$\frac{d^{\alpha}}{dx^{\alpha}}f(x)=\frac{d^{\alpha}}{dx^{\alpha}}\int_{-n}^{n}F(w)e^{2\pi ixw}dw=\int_{-n}^{n}F(w)\frac{d^{\alpha}}{dx^{\alpha}}(e^{2\pi ixw})dw=\int_{-n}^{n}F(w)(2\pi iw)^{\alpha}\left(e^{2\pi ixw}\right)dw$$

Eq. 4-34

$$\frac{d^{\alpha}}{dx^{\alpha}} f(x) = \int_{-\infty}^{\infty} F(w)(2\pi i w)^{\alpha} \left(e^{2\pi i x w}\right) dw = (2\pi)^{\alpha} (i)^{\alpha} \int_{-\infty}^{\infty} F(w)(w)^{\alpha} \left(e^{2\pi i x w}\right) dw \quad \text{Eq. 4-34}$$

$$\left|\frac{d^{\alpha}}{dx^{\alpha}} f(x)\right| \le \left|(2\pi)^{\alpha}\right| \int_{-\infty}^{\infty} F(w)|w|^{\alpha} dw \quad \text{Eq. 4-35}$$

$$\left|\frac{d^{\alpha}}{dx^{\alpha}} f(x)\right| \le \left|(2\pi)^{\alpha}\right| \int_{-\infty}^{\infty} \frac{C|w|^{\alpha}}{(1+|w|)^{\alpha+\beta}} dw \quad \text{Eq. 4-36}$$

$$\left|\frac{d^{\alpha}}{dx^{\alpha}} f(x)\right| \le K \int_{-\infty}^{\infty} \frac{|w|^{\alpha}}{(1+|w|)^{\alpha+\beta}} dw < \infty \text{ for } \alpha < \beta \text{ and } \beta > 1 \quad \text{Eq. 4-37}$$

As mentioned earlier in this section, a wavelet has n vanishing moments, if and only if for all positive integers $k < n$, it satisfies the equation

$$\int_{-\infty}^{\infty} x^{k-1} \psi(x) dx = 0. \quad \text{Eq. 4-38}$$

4.3 Wavelet Modulus Maxima

For any integer $p < n$, $F(w)$ can be factored into

$$F(w) = (iw)^{p} \psi^{1}(w) \quad \text{Eq. 4-39}$$

where

1. $F(w)$ is the Fourier Transform of $\psi(x)$ and
2. $\psi(x)$ is the p^{th} derivative of $\psi^{1}(x)$, i.e.

$$\psi(x) = \frac{d^{p}}{dx^{p}} \psi^{1}(x) \quad \text{Eq. 4-40}$$

and $\psi^{1}(x)$ is an admissible wavelet. Hence

$$WT\left[f(x)\right](a,b) = f(x) * \psi_{a}(x) = \frac{d^{p}}{dx^{p}} \left(f(x) * a^{p} \psi_{a}^{1}(x)\right) = a^{p} \left(\frac{d^{p} f(x)}{dx^{p}} * \psi_{a}^{1}(x)\right). \quad \text{Eq. 4-41}$$

The Wavelet Transform of $f(x)$ with respect to the wavelet $\psi(x)$ is thus equal to the Wavelet Transform of its p^{th} derivative, computed with the wavelet $\psi^1(x)$, and multiplied by a^p (a is the scale factor, and is inversely proportional to the frequency). If p is an integer such that $0 < \alpha - p < 1$ the function $f(x)$ is uniformly Lipschitz α on an open interval $]a,b[$ if and only if $\frac{d^p f}{dx^p}$ is uniformly Lipschitz $\alpha - p$ in the same interval.

It follows that the Wavelet Transform is well adapted to estimate the local regularity of functions.

The Wavelet Modulus Maxima is useful in locating all the singularities in the signal. This dissertation research makes use of this property of the wavelets to identify the singularities in the images of the brain to do registration. Mallat et. al. [Mallat89][mallat92.1][Mallat92.2] present the following argument regarding the Wavelet Modulus Maxima. If $\theta_a(x)$ is a smoothing function and $f(x)$ is some signal in $L_2(R)$, then edges with the scale a are local sharp variation points of $f(x)$ smoothed by $\theta_a(x)$. Consider wavelets $\psi^1(x)$ and $\psi^2(x)$ defined as

$$\psi_a{}^1(x) = \frac{d\theta_a(x)}{dx} \text{ and } \psi_a{}^2(x) = \frac{d^2\theta_a(x)}{dx^2} \qquad \text{Eq. 4-42}$$

$$WT^1[f(x)](a,b) = f * \psi^1_{a,b}(x) \text{ and } WT^2[f(x)](a,b) = f * \psi^2_{a,b}(x) \qquad \text{Eq. 4-43}$$

$$WT^1[f(x)](a,b) = f * \left(a\frac{d\theta_a(x)}{dx} \right) = a\frac{d}{dx}\left(f(x) * \theta_a(x)\right) \qquad \text{Eq. 4-44}$$

$$WT^2[f(x)](a,b) = f * \left(a\frac{d^2\theta_a(x)}{dx^2} \right) = a\frac{d^2}{dx^2}\left(f(x) * \theta_a(x)\right) \qquad \text{Eq. 4-45}$$

The Wavelet Transform $WT^1[f(x)](a,b)$ and $WT^2[f(x)](a,b)$ are proportional to the first and second derivative of $f(x)$ smoothed by $\theta_a(x)$. For a fixed scale a, $WT^1[f(x)](a,b)$ yields the local extrema with respect to the variable x and these correspond to the points where $WT^2[f(x)](a,b)$ changes sign, i.e, zero-crossings of $WT^2[f(x)](a,b)$. The locations where $WT^2[f(x)](a,b)$ changes sign, correspond to the inflection points of the function $f*\theta_a(x)$. The local extrema approach has several advantages over the zero-crossings approach. An inflection point can be either the maximum or minimum of the absolute value of the first derivative, see Figure 4-1. The local maxima of the absolute value of the first derivative are sharp variation points of $f*\theta_a(x)$, whereas the minima correspond to slow variation points. The two types of inflection points can be distinguished by observing the value of $WT^1[f(x)](a,b)$, but they cannot be differentiated using the location on the x-axis where $WT^2[f(x)](a,b)$ changes sign, i.e., from the zero crossings of $WT^2[f(x)](a,b)$. Therefore the singularities can be detected by observing the behavior of $WT^1[f(x)](a,b)$. The Canny edge detector uses a Gaussian smoothing function and is more suited for continuous functions. However, it does not have a natural built in scaling function. The Wavelet Transform has an advantage over the canny edge detector because it can be easily adapted to discrete signals, especially those using orthogonal wavelets. This will work well with image data. Another advantage of the Wavelet Transform Modulus Maxima method is that elegant recursive fast algorithms can be used to reduce computing time. We adapt the definition of S. Zhong [Zhong90] to define Wavelet Transform Modulus Maxima.

A modulus maxima occurs at a point (a_0, x_0), if (a_0, x_0) is a local maximum of the modulus, and furthermore in any neighborhood of (a_0, x_0) there is a point (a_0, x') with

$$|WT[f](a_0, x)| \leq |WT[f](a_0, x_0)|. \qquad \text{Eq. 4-46}$$

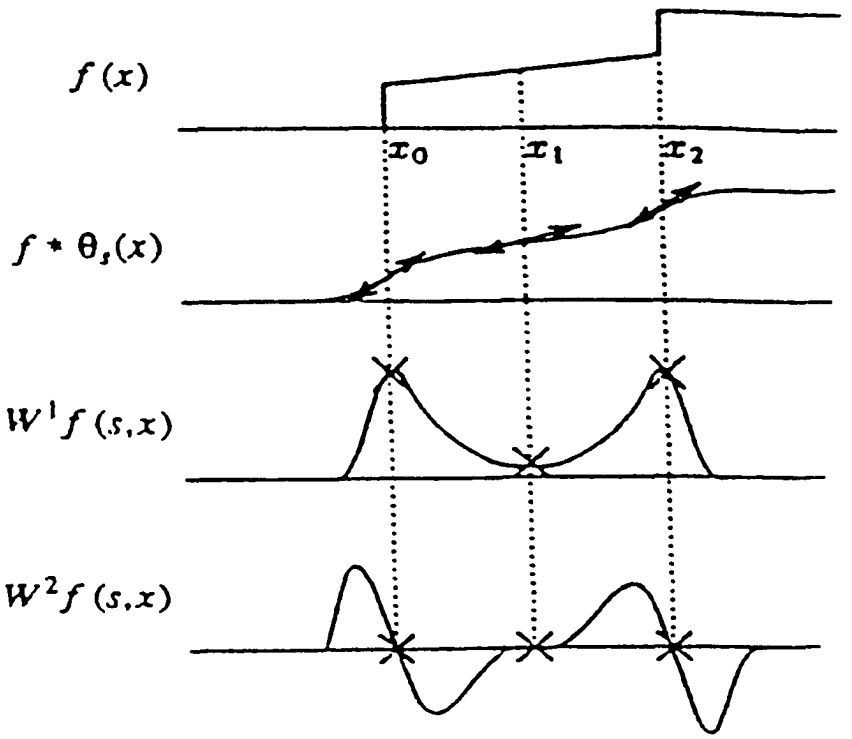


Figure 4-1 : This figure shows a function and then the plot of the function convolved with a smoothing function θ_a . The third plot from the top is a plot of the Wavelet Transform of the function, where the wavelet is the first derivative of the smoothing function. The last plot is again the Wavelet Transform of the function in the first plot. However the wavelet is so chosen that it is the second derivative of the smoothing function [Mallat92.1].

The local extremum is any point (a_0, x) where $\left(\frac{dWT[f](a_0, x_0)}{dx}\right)$ changes sign or has a zero-crossing at $x = x_0$, when x varies.

Using the derivation $|WT[f(x)](a,b)| \le Ca^{\alpha}$ and by taking the log of both sides of this equation

$$\log|WT[f(x)](a,b)| \le \log(C) + \alpha \log(a) \qquad \text{Eq. 4-47}$$

The Lipschitz regularity at x_0 is the maximum slope of straight lines that remain above $\log|WT[f(x)](a,b)|$, in log scale. For our purposes of registration, only the modulus maxima is used in detecting the singularities in the signal. We extend this definition to more than one dimension by considering the local maximum of the modulus for all the neighboring pixels.

Let $\theta(x,y)$ be a smoothing function. (By smoothing function we mean a function whose integral over x and y is equal to 1 and converges to 0 at "infinity"). For each scale let $\theta_a(x,y)$ represent the smoothing function at that a scale. Let $f(x,y)$ be a function representing the image gray values. Then the smoothed image (or low pass image) is obtained at different a scales by convolving $f(x,y)$ with $\theta_a(x,y)$. The gradient vector at any point is

$$\vec{\nabla}(f * \theta_a)(x,y) \qquad \text{Eq. 4-48}$$

The direction of the gradient vector at any point (x_0, y_0) indicates the direction in the image plane (x,y) along which the directional derivative of $f(x,y)$ has the largest absolute value. Edges are points where the modulus of the gradient vector is maximum

in the direction towards which the gradient vector points in the image plane. Let $\psi^1(x,y)$ and $\psi^2(x,y)$ be two wavelets where

$$\psi^1(x,y) = \frac{d\theta(x,y)}{dx} \qquad \text{Eq. 4-49}$$

$$\psi^2(x,y) = \frac{d\theta(x,y)}{dy} \qquad \text{Eq. 4-50}$$

Since we will be considering the Wavelet Transforms at different scales, we can represent the wavelet functions at different scales as:

$$\psi_a^1(x,y) = \frac{1}{a^2}\psi^1\left(\frac{x}{a},\frac{y}{a}\right) \text{ and } \psi_a^2(x,y) = \frac{1}{a^2}\psi^2\left(\frac{x}{a},\frac{y}{a}\right) \qquad \text{Eq. 4-51}$$

The Wavelet Transform can now be computed by considering the x direction and the y direction with components as shown :

$$WT^1[f(x)](a,b) = f(x,y) * \psi_a^1(x,y) \qquad \text{Eq. 4-52}$$

$$WT^2[f(x)](a,b) = f(x,y) * \psi_a^2(x,y) \qquad \text{Eq. 4-53}$$

The edge points can be located by computing the gradient vector at each point on the image plane using the two components $WT^1[f(x,y)](a,b)$ and $WT^2[f(x,y)](a,b)$. From an implementation stand point $WT^1[f(x,y)](a,b)$ represents the vertical edges and $WT^2[f(x,y)](a,b)$ represents the horizontal edges. The modulus of the gradient vector is proportional to

$$M_a f(x,y) = \sqrt{\left|WT^1[f(x,y)](a,b)\right|^2 + \left|WT^2[f(x,y)](a,b)\right|^2}, \qquad \text{Eq. 4-54}$$

where $M_a f(x,y)$ is the modulus of the gradient vector that can be computed at every point in the image plane at scale a. The angle θ_a of the gradient vector with the horizontal (x-direction) at scale a is given by

$$\theta_a = \tan^{-1}\left(\frac{WT^2[f(x)](a,b)}{WT^1[f(x)](a,b)}\right) \qquad \text{Eq. 4-55}$$

It should be noted that the modulus of the gradient vector and the direction is computed at every point in the image plane at scale a. The direction of the gradient vector, as mentioned earlier, at any point indicates the direction in the image plane, along which the directional derivative has the largest absolute value. From a practical standpoint, the Wavelet Modulus Maxima can also be computed by comparing the modulas of the Wavelet Transform at every internal point with its nearest eight neighbors to determine if the value is greater than one of these neighbors and greater than or equal to the rest of the neighbors. The location of the Modulus Maxima provides the edge information.

4.4 Wavelet computation

In the one-dimensional case the algorithm to compute the wavelet algorithm can be represented with the following diagram:

In this diagram the following notation is used:

- $f[n]$ is the input sequence.
- $L(Z)$ is the Z - Transform of the Low Pass Filter .
- $H(Z)$ is the Z - Transform of the High Pass Filter.
- $\downarrow 2$ implies down sampling by 2.

The input sampled sequence represented as $f[n]$ is subjected to a low pass filtering operation. The low pass corresponds to a smoothing filter function. The same image is also subjected to high pass filtering operation. The output from the high pass filtering operation provides the wavelet coefficients after down sampling of the coefficients. After

down sampling the output from the low pass filter, it is passed to the next stage where it is once again subjected to a low pass and high pass filter operation. The output from the high pass filtering after down sampling provides wavelet coefficients at that stage. The output from the low pass filtering operation after down sampling is passed on to the next stage. The output of the high pass provides details and high frequency information. The output from the low pass filtering operation provides a smoother image. The process is repeated to a stage, where the application environment dictates stopping the algorithm.

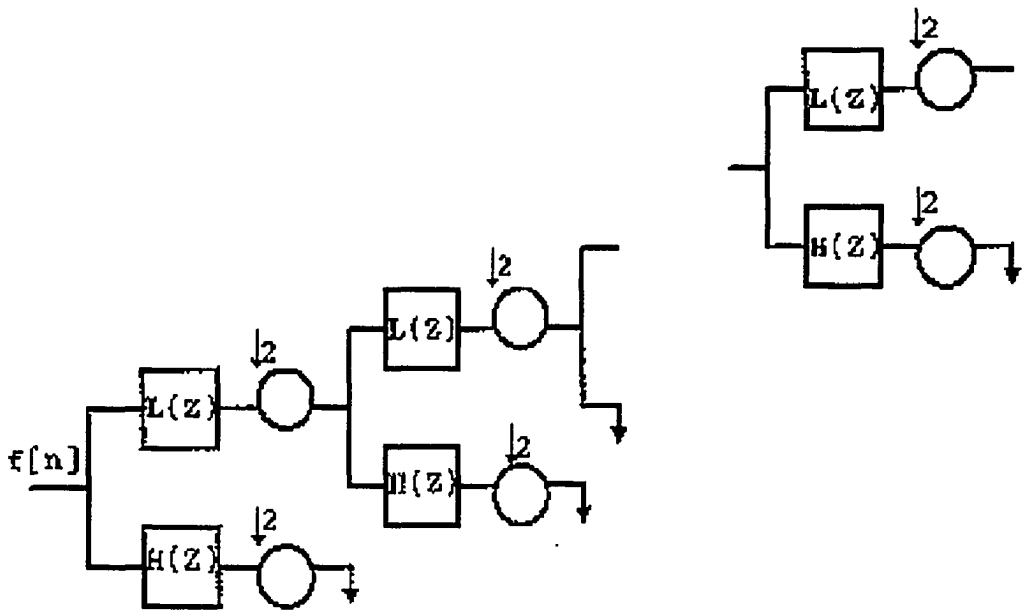


Figure 4-2 : Shows the decomposition scheme for wavelet analysis

As mentioned in the previous chapter, the filtering operation can be implemented as a multiplication in the frequency domain or convolution in the spatial domain. In the case of signals from images, the one dimensional case is generalized. We apply the Wavelet Transform to the rows and columns and to deal with the two dimensional signals, we assume the scaling function as separable. We then have :

$$\phi(x,y) = \phi(x)\phi(y)$$
$$\psi^1(x,y) = \psi(x)\phi(y)$$
$$\psi^2(x,y) = \phi(x)\psi(y)$$
$$\psi^3(x,y) = \psi(x)\psi(y)$$

Eq. 4-56

where

- $\phi(x)$ is the scaling function when we process in the horizontal direction (parallel to the x-axis)
- $\phi(y)$ is the scaling function when we process in the vertical direction (parallel to the y-axis)
- $\psi(x)$ is the wavelet that is used in the horizontal direction and
- $\psi(y)$ is the wavelet that is used in the vertical direction.

When an image is transformed from the spatial domain to the wavelet domain, we essentially subject each of the rows of the image, to low pass filtering with the scaling function, and then down sample them by two. This leads to an image matrix with N rows and $N/2$ columns. Each row of the original matrix is also subjected to high pass filtering and then down sampled by two. Once again we get a matrix with N rows and $N/2$ columns. So given an original image $f(x,y)$ shown in Figure 4-3, we obtain the image $f'(x,y)$ shown in Figure 4-4.

Each column of image $f'(x,y)$ is now subjected to a low pass filtering operation and then down sampled. The columns of the image $f'(x,y)$ are subjected to high pass filtering and down sampling. This results in four images each one a fourth the size of original image. In a typical Wavelet Transform coefficient computation, the image that is subjected to low pass in both the horizontal and vertical direction is taken as the input for

the second stage and this image is subjected to all the operations the original image was subjected . This process is repeated until diminishing returns from the application environment are met stopping the process.

Original Image $f(x,y)$ of an MRI of the brain

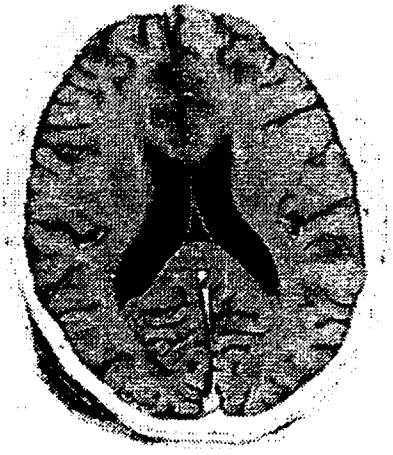


Figure 4-3 : MRI scan of the brain.

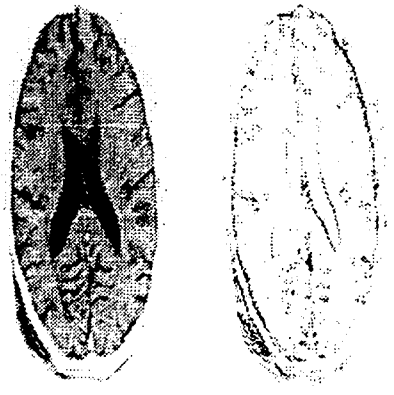


Figure 4-4 : Partial Analysis of the MRI of the brain.

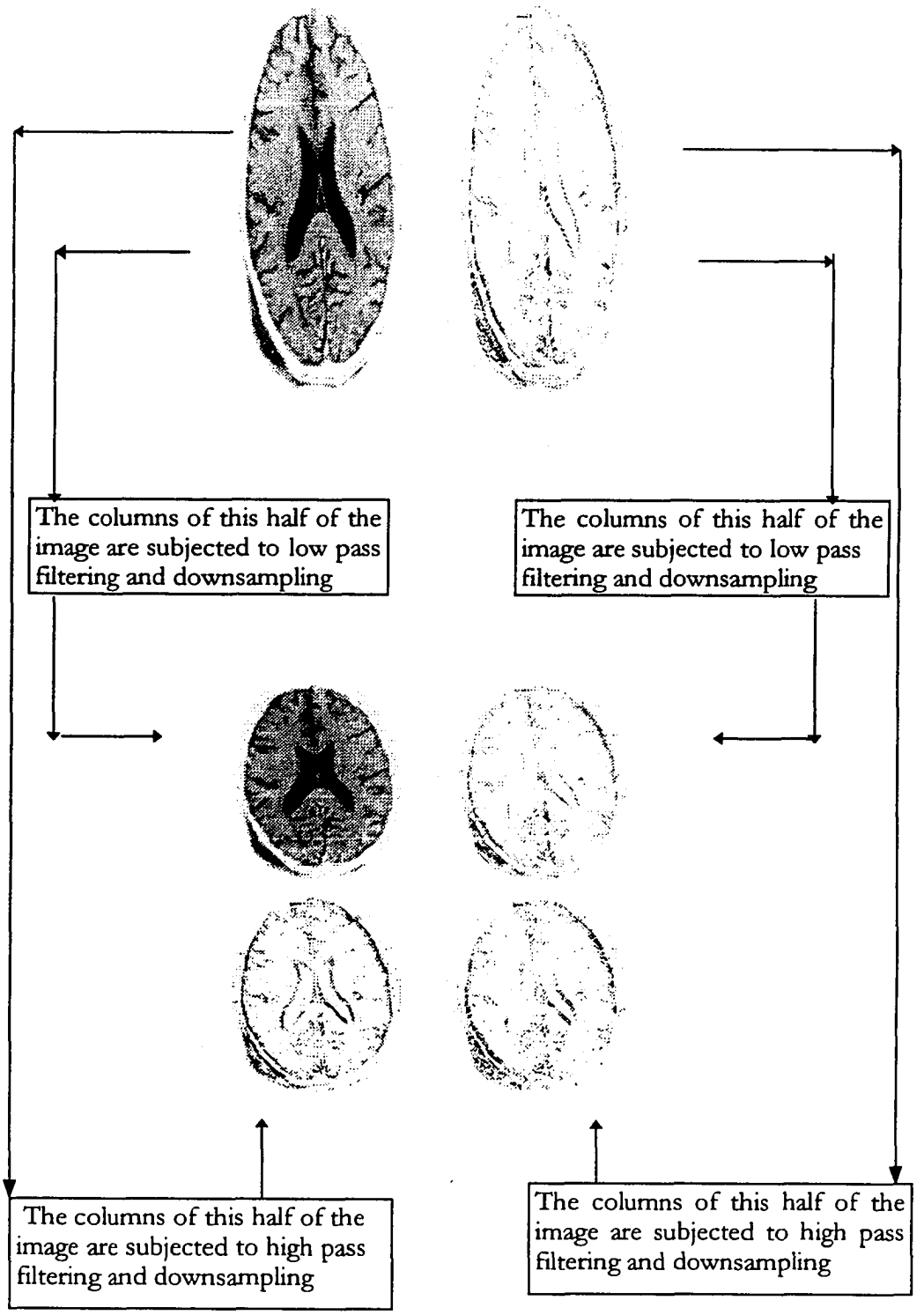


Figure 4-5 : Wavelet decomposition of a brain image

At each stage an input image $f(x,y)$ of size $N \times N$ is decomposed into four sub-images of size $(N / scale) \times (N / scale)$:

- An image is smoothed in the x and y direction; obtained by convolving $f_{2^j}(x,y) * \phi(x,y)$
- An image that has been subjected to high pass filtering in the horizontal direction can contain information about vertical edges; obtained by the convolving $f_{2^j}(x,y) * \psi^1(x,y)$.
- An image that has been subjected to high pass filtering in the vertical direction can contain information about horizontal edges; obtained by convolving $f_{2^j}(x,y) * \psi^2(x,y)$.
- An image that has been subjected to high pass filtering in the horizontal and vertical direction can contain information about diagonal edges; obtained by convolving $f_{2^j}(x,y) * \psi^3(x,y)$.

The image that was subjected to low pass filtering operations serves as the input image for the next stage. Figure 4-5 shows the MRI of the brain during and after the first stage of the analysis. Figures 4-6 and 4-7 show the next two stages in the process of performing the Wavelet Transform on the MRI image.

Using the image containing horizontal edges (Figure 4-8) and vertical edges (Figure 4-9), the the Wavelet Modulus image and the Wavelet Modulus Maxima image are computed. Figure 4-10 shows the modulus image after the first stage. Figure 4-11 shows the modulus maxima image after stage 1. The iterations can be repeated for several stages depending on the demands of the application domain.

4.5 Conclusion

In order to register images, the computation of the Wavelet Modulus Maxima Image is one of the key steps.

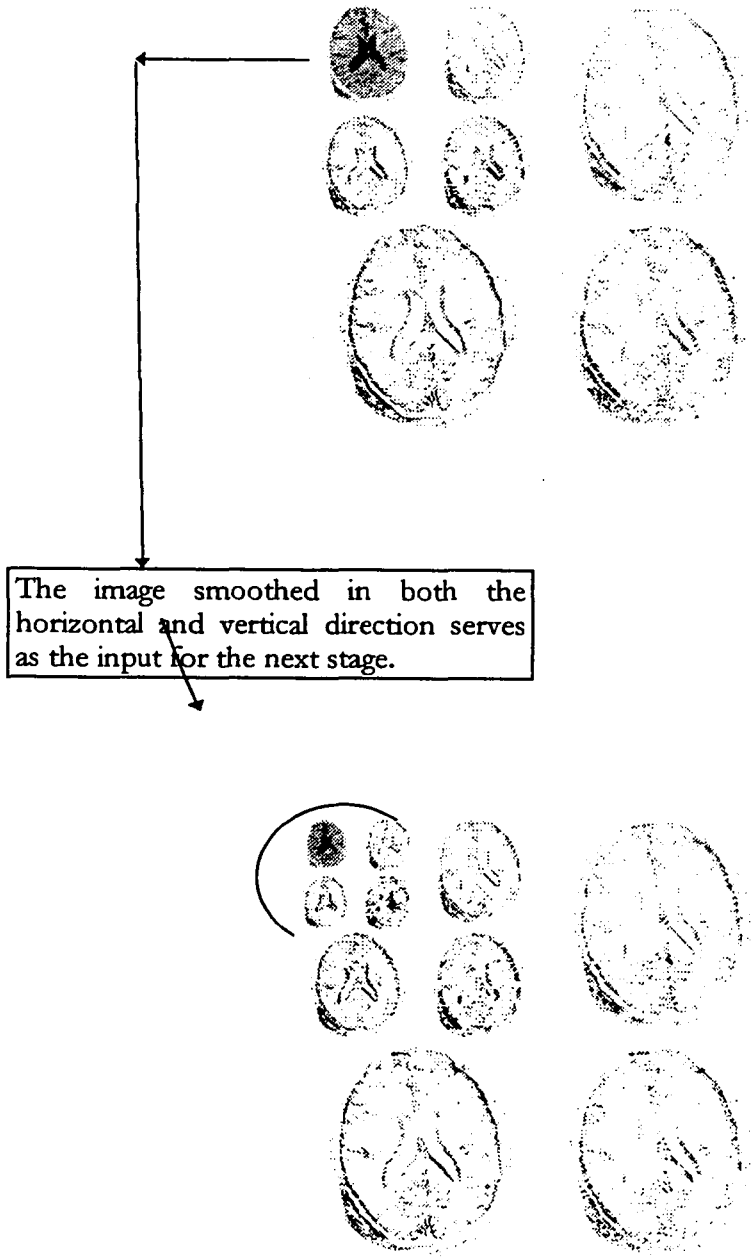


Figure 4-7: Stage three of wavelet decomposition of an MRI image of the brain

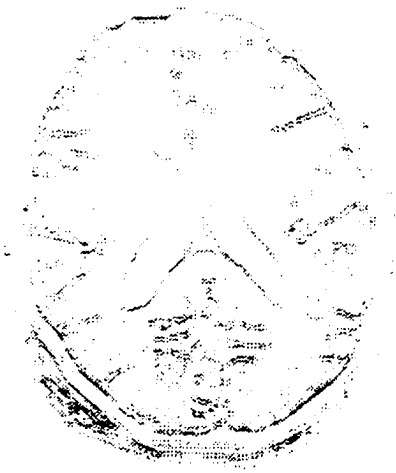


Figure 4-8 : Image showing horizontal edges

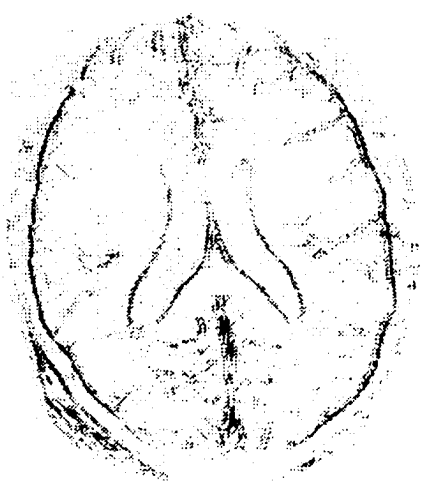


Figure 4-9 : Image showing vertical Edges

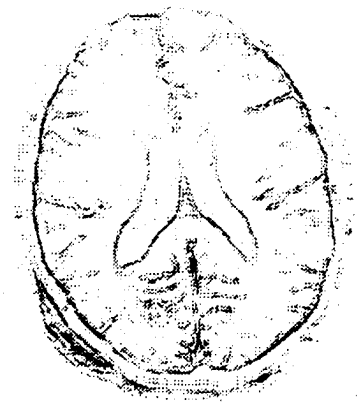


Figure 4-10 : Shows the Modulus Image

Figure 4-11 : Image showing the Wavelet Modulus Maxima Image

This chapter dealt with the computation of the Wavelet Modulus Maxima image and the computation of the Wavelet Transform. Both theoretical and practical issues were explored, however, several implementation issues need clarification. The next chapter addresses implementation of these issues as well as our methodology to register images.

CHAPTER 5

REGISTRATION, IMPLEMENTATION AND VALIDATION

5.1 Introduction

This chapter details the step by step procedure to register images and implementation issues relating to the registration procedure. A section of this chapter is devoted to issues related to the validation of registration techniques used in this dissertation.

5.2 Steps in the registration process

In this section a step by step procedure to register two images is presented.

Step 1 - Compute the Wavelet Modulus Maxima for both images

When we register one image with another we determine Wavelet Modulus Maxima for each image. In Chapter 4 the procedure to obtain the Wavelet Modulus Maxima of any image was presented. For this discussion, we consider the two simple images shown in Figures 5-1 and 5-2. These are MRI images of the brain obtained from the same patient that have been rotated and shifted. The Wavelet Modulus Maxima for the images in Figures 5-1 and 5-2 are shown in Figures 5-3 and 5-4 respectively.

Step 2 - Find the convex hull for both the Wavelet Modulus Maxima Images

We use the Akl-Tousant Hull Algorithm to determine the points on the Convex Hull. This algorithm is included in Appendix C of this dissertation. The Wavelet Modulus Maxima Image obtained in Step # 1 is the input for the Akl-Tousant Hull Algorithm.

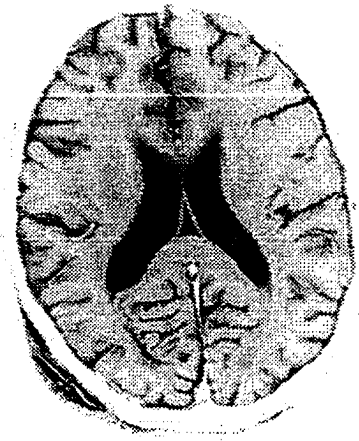


Figure 5-1 : Image # 1 MRI of the brain

Figure 5-2 : Image # 2 (Image # 1 rotated)

We take the original image and obtain the Wavelet Modulus Maxima Image

Figure5-3 : Modulus Maxima Image for image on Figure 5-1

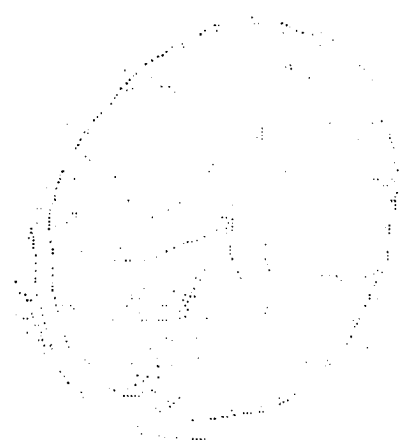


Figure 5-4 : Modulus Maxima Image for image on Figure 5-2

The convex hull program uses the image displayed in Figure 5-3 as input and determines the points on the convex hull displayed in Figure 5-5. Similarly using the image displayed in Figure 5-4 as input the hull program produces the set of points on the convex hull displayed in Figure 5-6.

Figure 5-5 : Hull points obtained using the image on Figure 5-3.

Figure 5-6 : Hull points obtained using the image on Figure 5-4.

Step 3 - Move each image to the center of the image frame.

Using the points that form the convex hull for each image, find the centroid for each image from of the points on the hull. Then move the original image using the centroid to the center of the image frame. Figures 5-7 and 5-8 show the original images centered in the image frame.

Step 4 - Find the principal components of both images and compute the angle.

Find the principal components of the images using the points on the hull. Using the principal axis of these points, we compute the rotation angle. We then rotate one of the two images by this rotation angle. Since both images were moved to the center of the frame, this movement is actually a translation. After this translation

followed by rotation of one of the images, the two images are registered. Figure 5-9 shows the rotated image. The registered images can be fused in a variety of ways. The fusion of registered images is a separate important area that requires image registration. Each application dictates how images are fused. In Medical Imaging, the modalities can dictate the fusing technique.

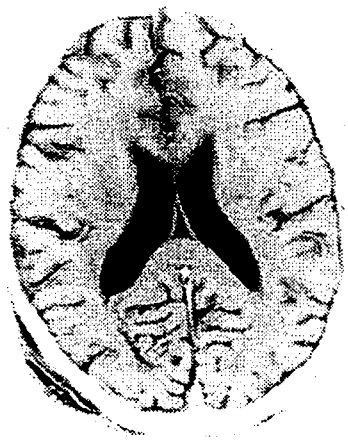


Figure 5-7 : Original Image

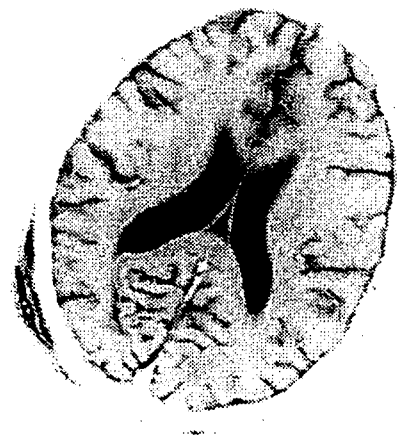


Figure 5-8 : Rotated Image

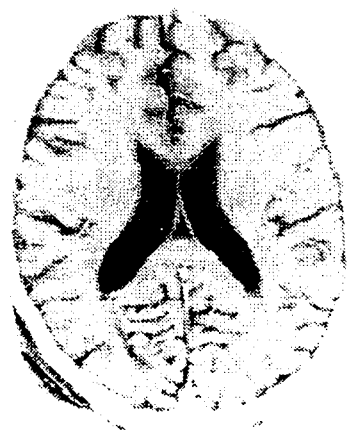


Figure 5-9 : Registered Image

Images in Figures 5-7 and 5-9 are registered using the four steps given in this section. Even though four different steps were used for the explanation of the method, software can be written to integrate these steps into one. The next chapter provides a pseudo code for some of the steps, in order to provide better clarity from a programming point of view. Iteration is used to improve the accuracy.

5.3 Implementation Issues

This section deals with implementation issues about processes, software and programming.

5.3.1 Software

The software was written in a C++ environment, however the software could have been written in any other language. Though the programs were compiled in a Visual C++ environment only C features were used. Two reasons prompted this choice:

1. The Visual C++ provided type checking and better adherence to a safer programming environment. By using the C++ compiler, this was ensured in the programming process; however the object oriented features were not used.
2. The C language is widely available especially in embedded system environments. Thus a greater portability of this code is achieved.

The programming was actually done on a PC; however the code can be easily adapted to other environments like UNIX, etc.

Particular attention was paid to features like order of multiplication, use of double and long double where precision was important. Software engineering principles calls for partitioning the code into functions. This was done only as far as efficiency was not sacrificed. Macros were used to provide readability as well as for inlining the code.

5.3.2 Wavelet Computation

5.3.2.1 Signal Extensions

The first issue that the software addresses is extension of the signal. This is necessary because we are dealing with finite signals when we use signals from images. In the process of computing the Wavelet Transform, we must convolve the input signal (image) with the high pass and low pass filters at each stage. This convolution causes border distortions to the signal in all finite signals. When this occurs there are three options available to extend the signal in the time domain leaving the gray code field to resolve this issue:

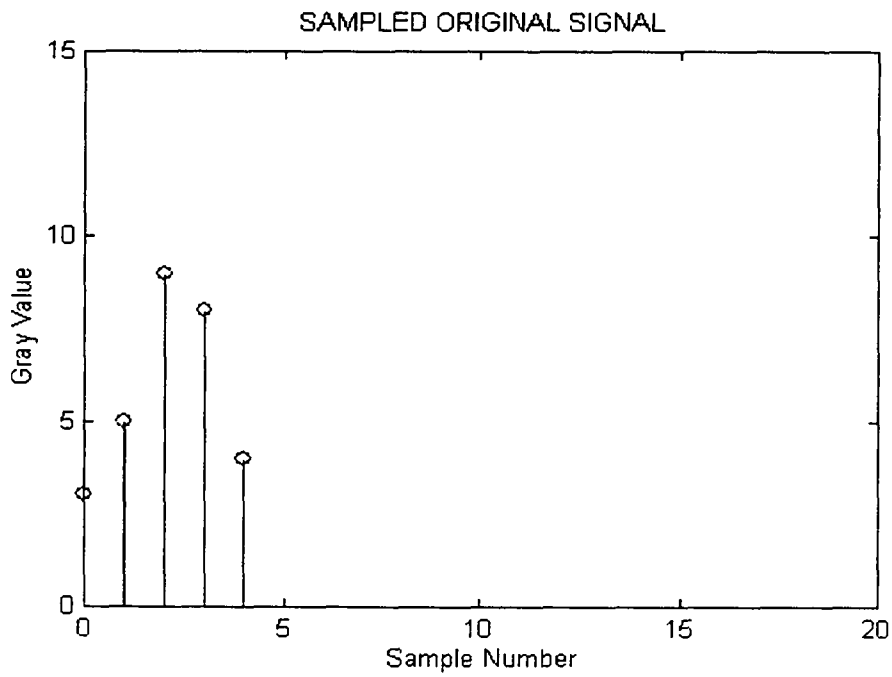


Figure 5-10 : Original sampled signal

- Extending the signal by zeroes both before and after the actual signal. Figure 5-10 shows an example of a one dimensional signal and Figure 5-11 shows the signal padded with zeros shown as dots. In signal theory, extending signals by zeros is

referred to as Zero-padding. Zero-padding introduces a jump in the function and causes border distortions in normal images that extend to the frame boundary.

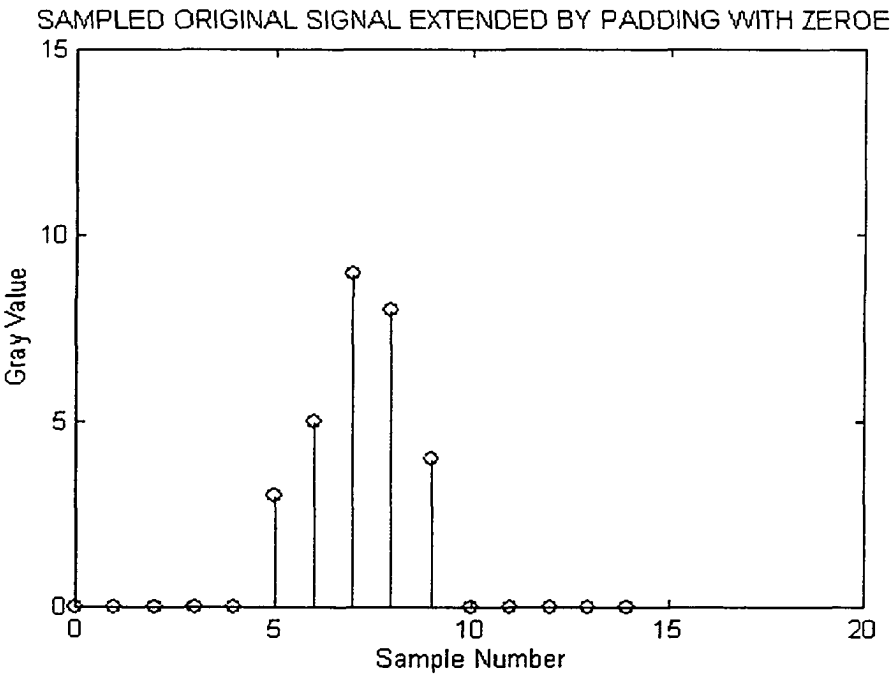


Figure 5-11 : Original sampled signal with zero padding

- The signal is extended by repeating the signal. This is good if the signal is genuinely periodic or almost periodic. The filter matrix becomes circulant and can be implemented by using the modulus operation on the index of the matrix holding the filter. However it also produces a jump in the signal introducing border distortions. Figure 5-12 shows the signal in Figure 5-10 extended by periodicity. In signal processing, when a signal is extended by repeating the signal, it is often referred to as "extended by periodicity".
- Signals can be extended by reflecting the signal at either end. This is referred to in signal processing language as symmetric extension. Symmetric extension extends the images in a continuous way at the border.

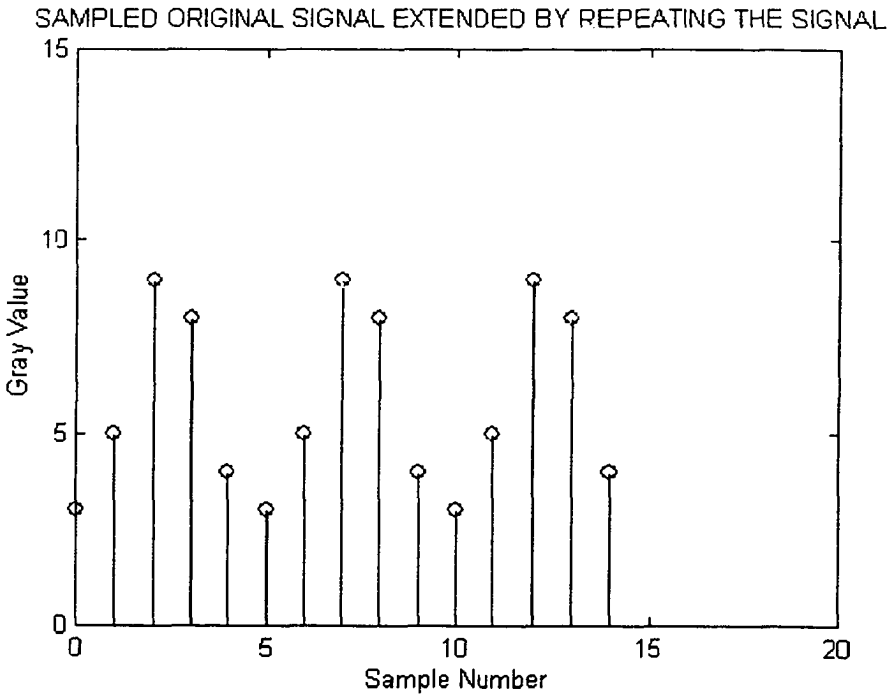


Figure 5-12 : Original sampled signal with periodic extension

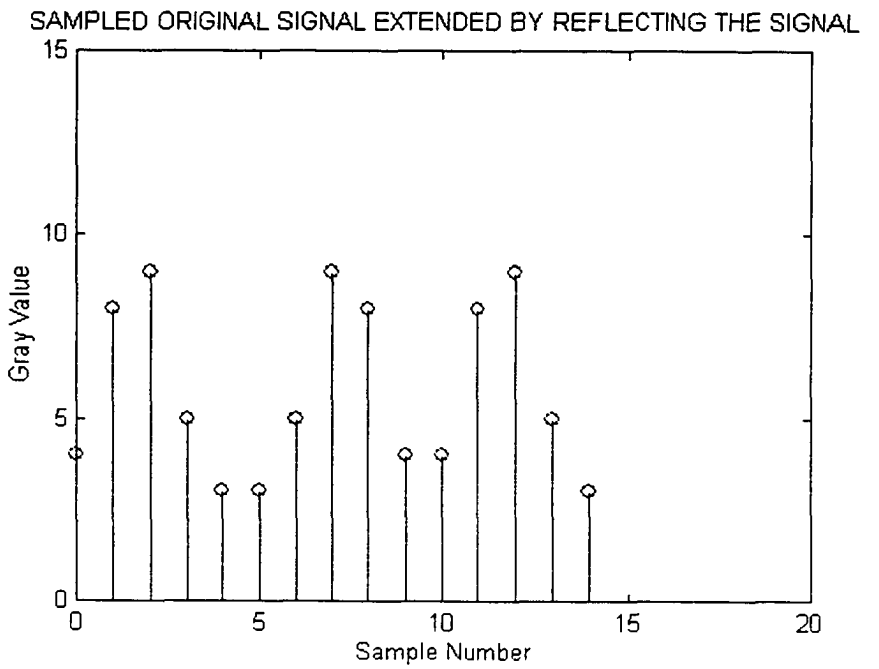


Figure 5-13 : Original sampled signal with symmetric extension

Although the objective performance of circular convolution is good, symmetric extension produces the best results according to image processing literature [Strang96][Proakis95]. Figure 5-13 shows the original signal in Figure 5-10 extended using the technique of symmetric extension. In our use of MRI images of the brain, the brain image often resided several pixels away from the border and the choice of extension had no negative effect. But since this registration process could be generalized for use with other application areas, symmetric border extensions were chosen and implemented.

5.3.2.2 Choice of Wavelets

Wavelet theory and its application to image processing provides us with the knowledge that orthogonal wavelets are not symmetric. However, biorthogonal wavelets are symmetric. Additionally, we wanted the filters to be maximally flat half band filters. Maximally flat filters are filters that have more zeroes (z-domain) at frequency $\omega = 0$ and $\omega = \pi$. If we chose filters that are not maximally flat, they tend to produce a checkerboard effect according to Strang and Nguyen[Strang96]. In the spatial domain this means that we are interested in an increasing number of vanishing moments being zero. Increasing the number of vanishing moments produces smoother wavelets with more oscillations. The consequence of producing smoother wavelets is that the number of filter coefficients increase [Strang96]. As the number of filter coefficients increase, the computational efficiency is adversely affected. So, the choice of wavelets is a tradeoff between these competing interests. In the registration process, we are primarily interested in the analysis phase of the Wavelet Transform and not in synthesis or reconstruction, and in this strict sense, these considerations regarding a checkerboard effect, ringing, etc. are not an important. However, it is likely that wavelets will become a larger part of Medical Imaging in the future and registration may serve as a necessary preprocessing step. If that

happens, number of coefficients which impacts the computational efficiency becomes more important in registration of images. Hence, this factor was taken into account in our choice of wavelets. Even so, a large number of candidate wavelets could have been selected for our purpose of computing the Wavelet Modulus Maxima because they satisfied the needed conditions.

For the purposes of this dissertation we selected and used the quadratic spline wavelet shown in Figure 5-14. It is antisymmetrical, and has small compact support. The cubic spline serves as the smoothing (scaling) function. The derivative of the quadratic spline is the cubic spline. Figure 5-15 shows a plot of the cubic spline function. Haar Wavelets are the simplest of the wavelets and the scaling function is simply a box function. Haar Wavelets have one vanishing moment, which means that they do not approximate functions very rapidly. Similarly, the wavelet coefficients when Haar Wavelets are used do not tend to zero as fast as almost all other known wavelets, at finer levels. Therefore, Haar Wavelets do not produce as much contrast in coefficient size between smooth and non-smooth sections of the image as wavelets with more vanishing moments. Daubechies Wavelets, the Morlet Wavelet, the Mexican Hat Wavelet, the Coiflet Wavelets and the Symlet Wavelets do not satisfy the condition that the wavelet function is the derivative of the scaling function. This is a necessary condition that the wavelet must satisfy in order for the Wavelet Modulus Maxima technique, according to Mallat [Mallat89] [Mallat92]. The next chapter, presents results of using this registration process on a number of data sets. There are other considerations that serve as determinants for the choice of filters used in any application where wavelets are used. Signals processing using filters, require us to permit either low frequency or high frequency or other frequency components of the signal to pass through the filter but then may require the attenuation of the frequency

components that are to be filtered out of the processed signal. Computation of the Wavelet Transform coefficients require subjecting the signal to high pass and low pass filters during each phase of the computation, as explained in the previous chapter. Ideally in all such cases we would like to have a perfect stop band attenuation. However, there is no ideal filter that provides a perfect stop band attenuation, when finite impulse response filters are considered. We encounter, the well documented phenomenon known as the Gibbs phenomenon[Rasmussen93] causing frequency leaks and stop band errors. Conventional digital signal processing techniques settle for the stop band error to be equiripple [Strang96][Proakis95][Oppenhein89].

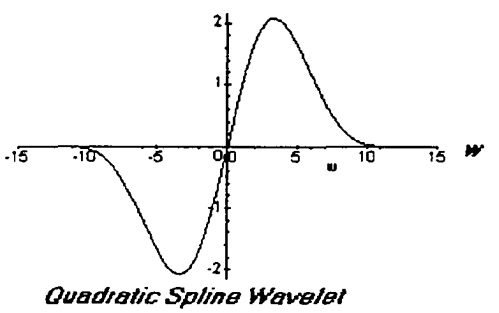


Figure 5-14 : Quadratic Spline Wavelet

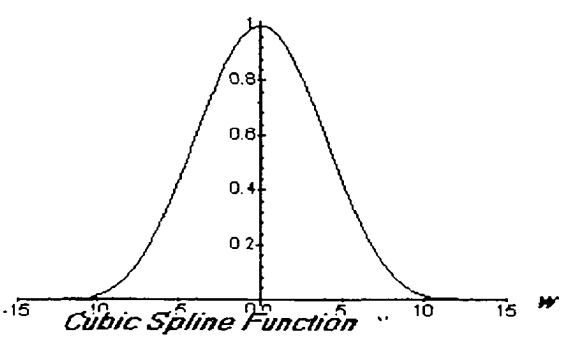


Figure 5-15 : Cubic Spline Function. (Derivative of the Quadratic Spline Wavelet)

5.3.3 Convex Hull Algorithm

There are several algorithms that can be used to find the convex hull of a set of points. We used the algorithm by Akl and Toussaint [Akl78] [Akl79]. This algorithm is presented in Appendix A. The algorithm has computational time complexity $O(n\log n)$. This is a fast algorithm that can be easily parallelized. We chose this algorithm from among the many available and equally efficient algorithms because of its simplicity. The data structure used is a doubly linked list which allows back tracking and elimination of points not on the hull. However if the dimensionality of the image were to be increased, this algorithm must to be modified or replaced.

5.3.4 Principal Component Analysis

This program is designed and written for two dimensional images. However this program can be easily modified for images with more than two dimensions. The code for this program computes the eigenvalues and eigenvectors for both images plus the angle between the principal axes of the two images. The details of this computation are presented in Appendix C.

5.3.5 Translation

Each image of the brain is "moved" to the center of the image frame. The centroid of each image is computed using the Wavelet Modulus Maxima Image. Based on the coordinates of each centroid, the required displacement is determined. The center of the image frame is located at the coordinate location [IMAGESIZE/2, IMAGESIZE/2], where IMAGESIZE is the number of pixels in one dimension of the image. For a 256×256 pixel image, the IMAGESIZE is 256. Sometimes, when an image is translated, parts of the image fall outside the new image frame. When this occurs, the part of the

image outside the frame is discarded. It is assumed that the parts of the image that are outside of the convex hull constitute no actual part of the MRI of the brain. It is our assumption that no useful information is lost. Figure 5-16 reflects this part of the implementation.

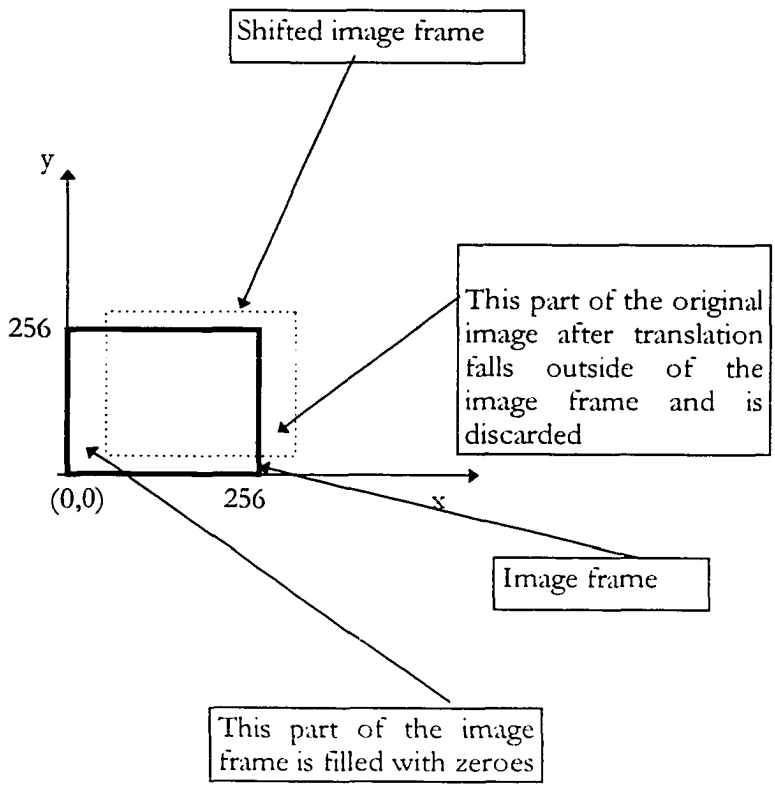


Figure 5-16 : Shows the effects of translation of the image frame.

5.3.6 Rotation

The rotation has its origin of the frame of reference at the center of the image frame. The rotation matrix for clockwise rotation through angle θ is given by

$$R = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \cos\theta \begin{bmatrix} 1 & k \\ -k & 1 \end{bmatrix} \qquad \text{Eq. 5-1}$$

where $k = \tan\theta$. Implementing the matrix with k saves extra multiplication [Strang96]. When we rotate the image frame about the center of the frame, the corners of the rotated image frame are no longer contained within the unrotated frame. We discard the parts of the image that are no longer contained within an unrotated frame as shown in Figure 5-17.

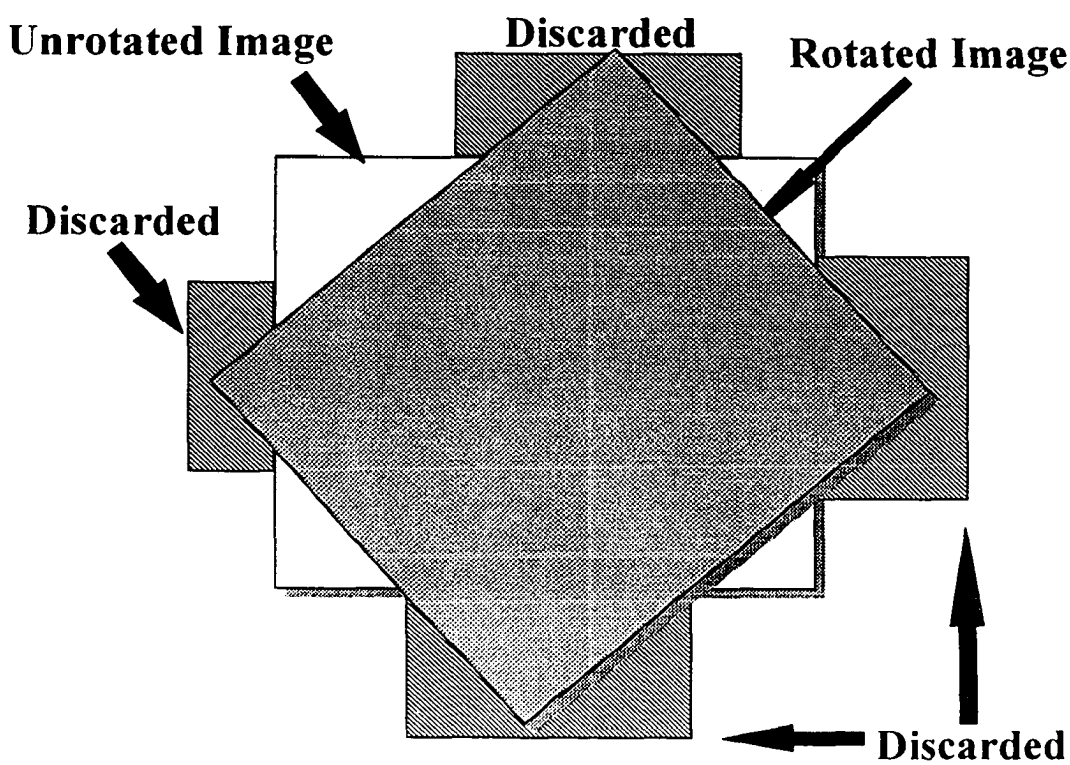


Figure 5-17 : Shows the effect of rotation of the image frame

5.4 Validation

The question to be answered is: How does one know if the objects have been registered correctly? In medical imaging, there is no unique or standard method to validate the registration process. In current professional meetings on Image Registration, this issue has been vigorously discussed and debated. There are several practices that are normally used to gauge the accuracy of registration which will be examined.

5.4.1 Using a physician to validate

In many registrations, physicians identify the points on objects that are to be registered. Since they are responsible for the diagnosis, they act as the judge whether the registration is accurate. Studies have shown that this method is very physician dependent and hence subjective. For lack of a standardized method for registration, this method of validation is the usual practice.

5.4.2 Using the difference Image to validate

At the outset, when this dissertation was proposed, it was believed that the difference image would serve as a measure to judge the accuracy of registration. The difference image would have longer run lengths once the images are registered. At first glance, this appears to be an excellent idea. However, run length encoding has some intrinsic problems in the way it encodes data. The run length encoding algorithm is only efficient for images that have long run lengths. However, if images have short lengths, run length encoding can double the size of the encoded image. Our experimental results produced almost doubled encoded image sizes. For example an MRI of the brain in PGM format has an image size of 65,607 bytes and its run length encoded size balloons to116,048 bytes, which is almost double the size of the original image. Run length encoding has an analogous effect on its difference image file. According to Castleman [Castleman96] difference images produce

strong gradients. If a registration process is not accurate, the difference image is one which has strong gradients. The method we have proposed determines the translation and rotational parameter accurately and produces resultant registrations with an accuracy within two degrees of perfect registrations in the tests used.

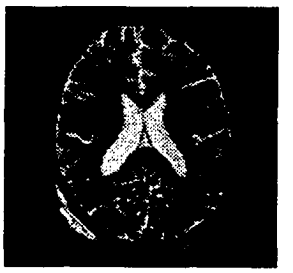
5.4.3 Using a Scatter plot to validate

The scatter plot is an excellent tool to verify if two images are correlated. The scatter plot of two images $f_1(x,y)$ and $f_2(x,y)$ is obtained by plotting the points (G_1, G_2) where G_1 is the gray level at any pixel position (x, y) in the first image and G_2 is the gray level at the same pixel position in second image. We obtain all the points in the scatter plot by considering every pixel location in both images. If two images are completely correlated, then these points will cluster about the 45^0 degree line passing through the origin. Our experimental results show that this happens only if the two images are perfectly registered and similar.

It is sometimes difficult to notice translation or rotation in normal video, hence the MRI images of the brain is printed in reverse video format. In reverse video, gray level of zero is represented by black and gray level of 255 is represented by white.

Figure 5-18 shows a plot of an image of LENA with an MRI. We see that the scatter plot is well spread out. However if we look closely at the scatter plot of the two brain MRI's in Figure 5-19 most of the points tend towards a 45^9 line passing through the origin. However as images are aligned spatially we see that the scatter plot begins to show points falling more or less on a 45^9 through the origin. To demonstrate the effect of spatial misalignment due to rotation, on the scatter plot and how scatter plots can be used in registration, Figures 5-20 through 5-28 were developed.

IMAGE # 1 (LENA) 

IMAGE # 2 (MRI of the brain) 

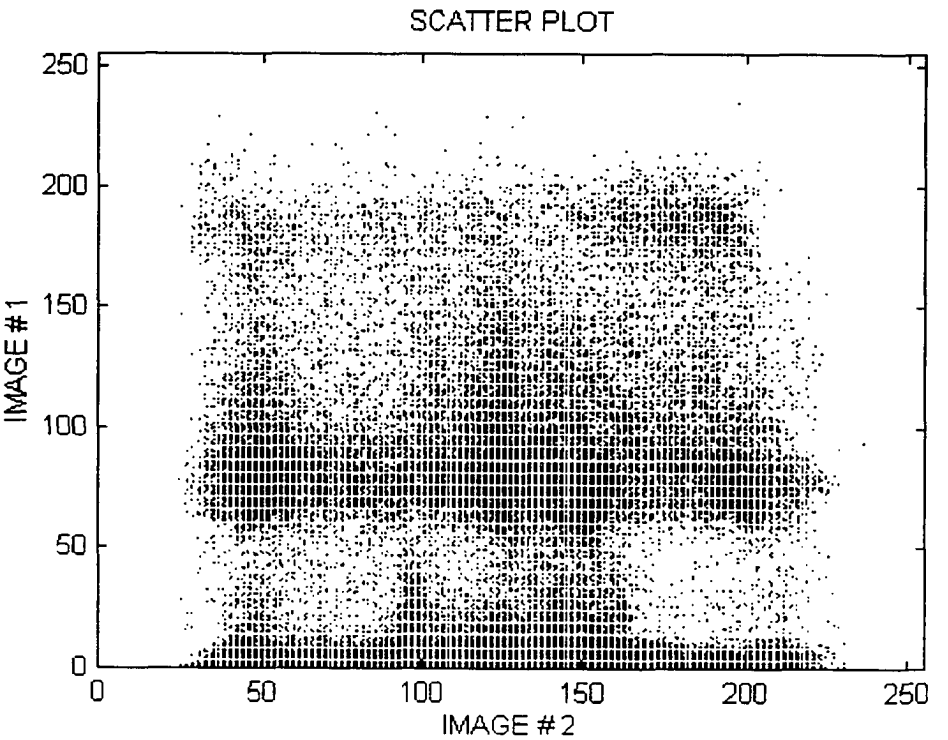


Figure 5-18 shows the scatter plot of two uncorrelated (LENA and MRI shown) images.

There is no clustering observed about a 45° line passing through the origin.

If two images are similar, the scatter plot shows some clustering about the 45^0 line through the origin. Figure 5-19 shows the scatter plot of two MRI images of the brain. The images are similar in the sense that they are MRI images of the brain, but come from two different patients.

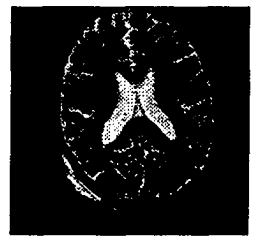
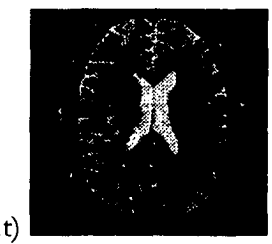
IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (MRI of the brain obtained from a different patient) 

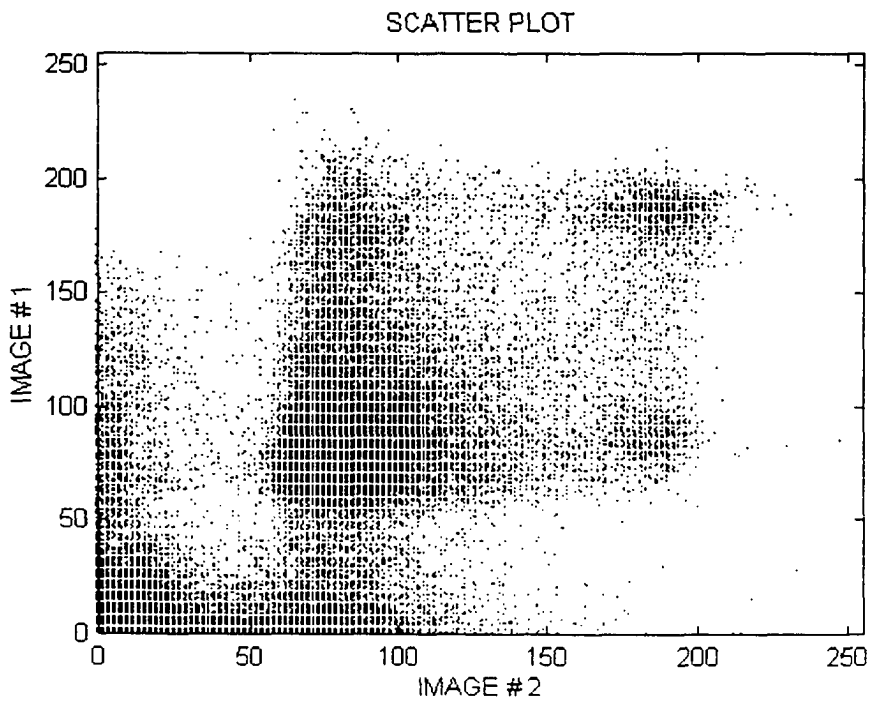


Figure 5-19 : Scatter plot of brain images obtained from two different patients.

Figure 5-20 shows the scatter plot of a brain image with a rotated version of the same image where the rotation is by 25^0.

IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Image #1 rotated by 25^0) 

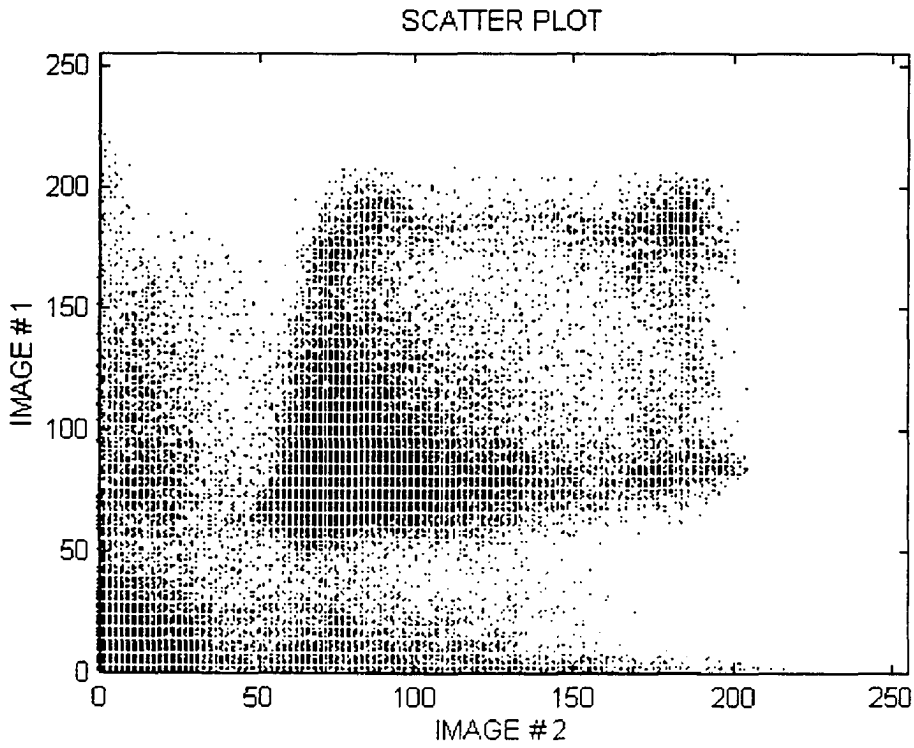


Figure 5-20 : Scatter plot of an MRI brain image with its rotated version (25^0).

Figure 5-21 shows the scatter plot of a brain image with a rotated version of the same image where the rotation is by 20^0.

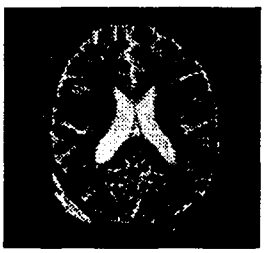
IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Image #1 rotated by 20^0) 

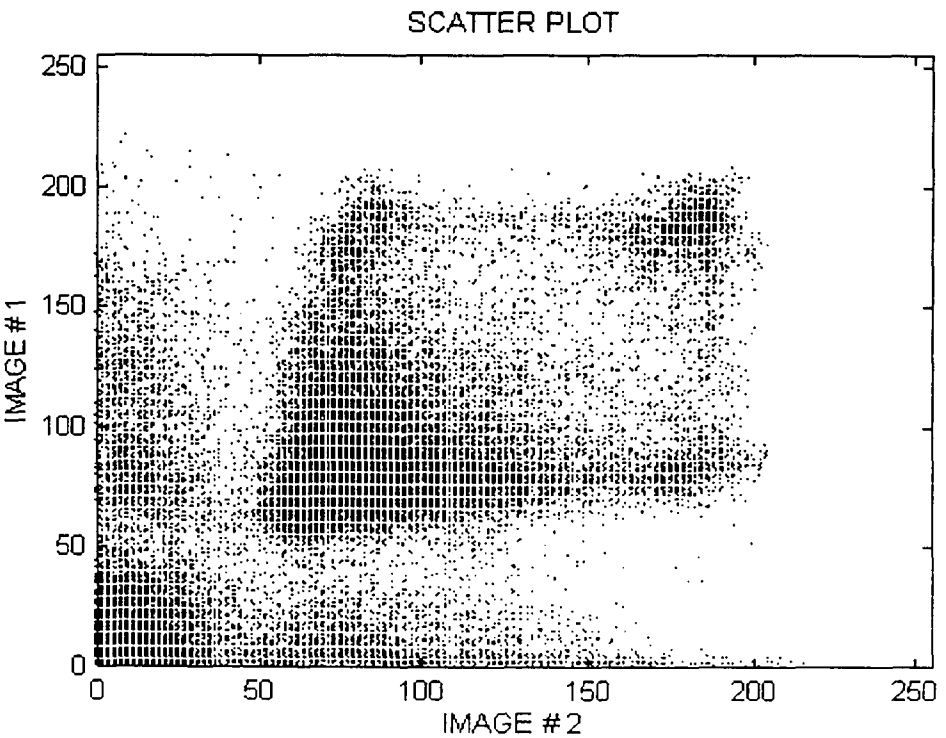


Figure 5-21 : Scatter plot of an MRI brain image with its rotated version (20^0).

Figure 5-22 shows the scatter plot of a brain image with a rotated version of the same image where the rotation is by 10^0.

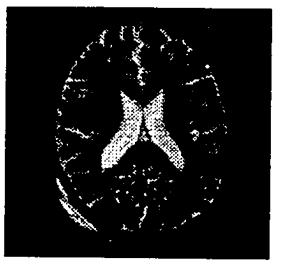
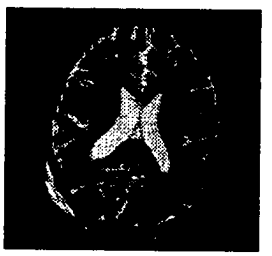
IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Image #1 rotated by 10^0) 

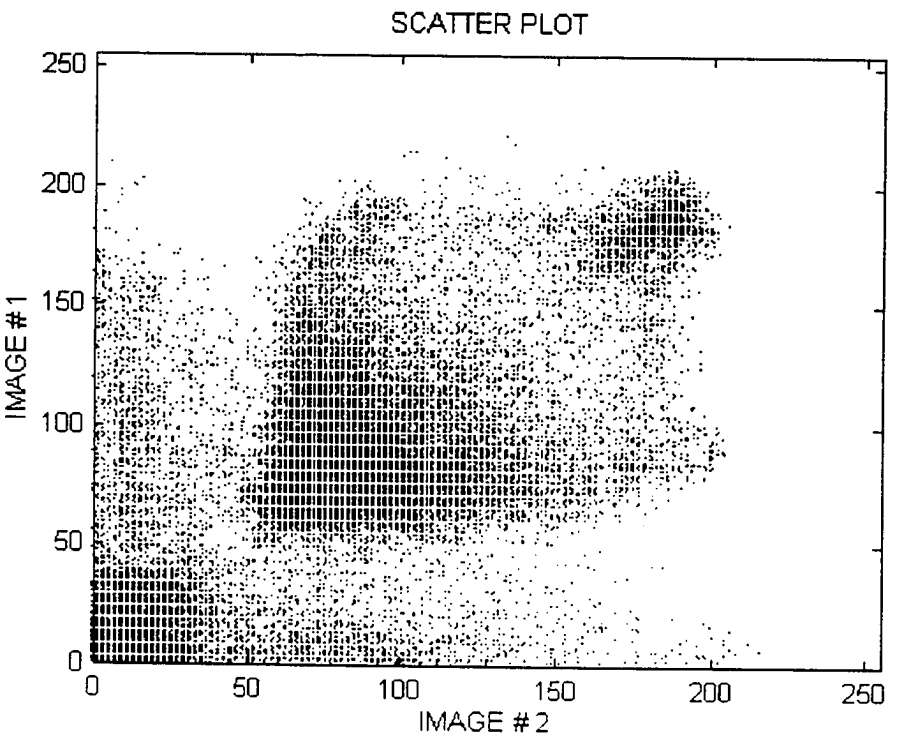


Figure 5-22 : Scatter plot of an MRI brain image with its rotated version (10^0).

Figure 5-23 shows the scatter plot of a brain image with a rotated version of the same image where the rotation is by 5^0.

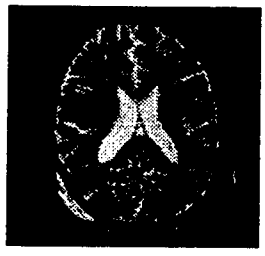
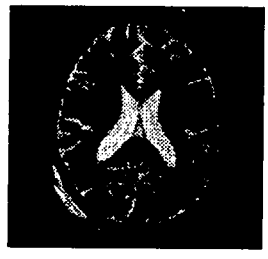
IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Image #1 rotated by 5^0) 

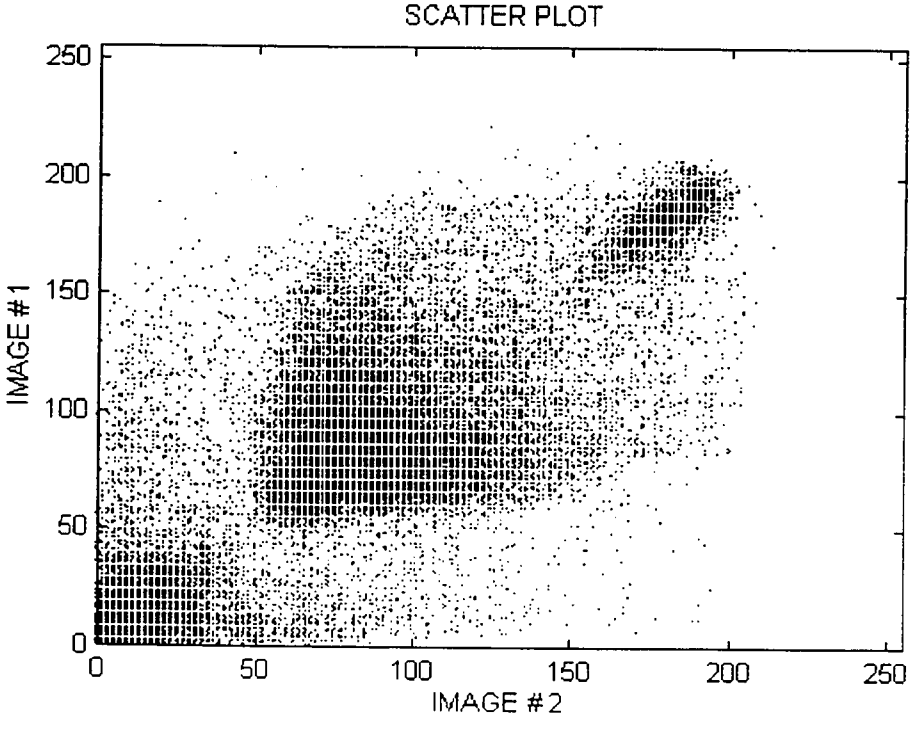


Figure 5-23 : Scatter plot of an MRI brain image with its rotated version (5^0).

Figure 5-24 shows the scatter plot of a brain image with a rotated version of the same image where the rotation is by 4^0.

IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Image #1 rotated by 4^0) 

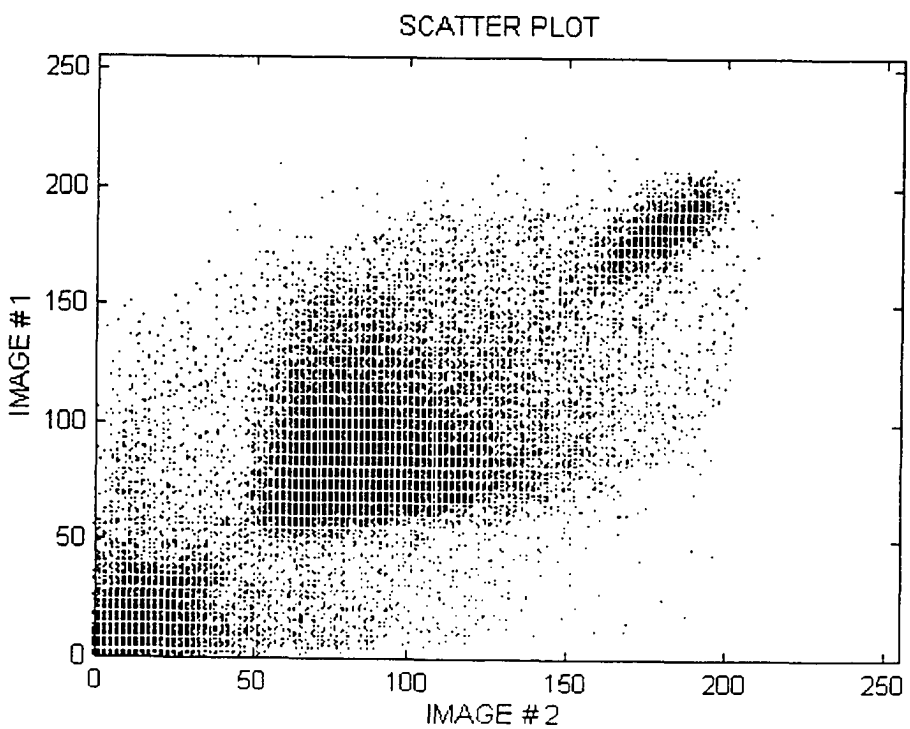


Figure 5-24 : Scatter plot of an MRI brain image with its rotated version (4^0).

Figure 5-25 shows the scatter plot of a brain image with a rotated version of the same image where the rotation is by 3^0.

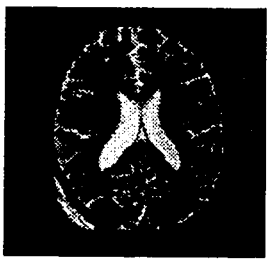
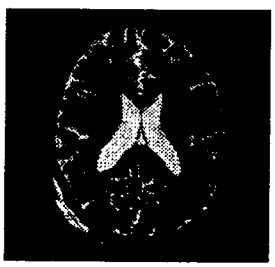
IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Image #1 rotated by 3^0) 

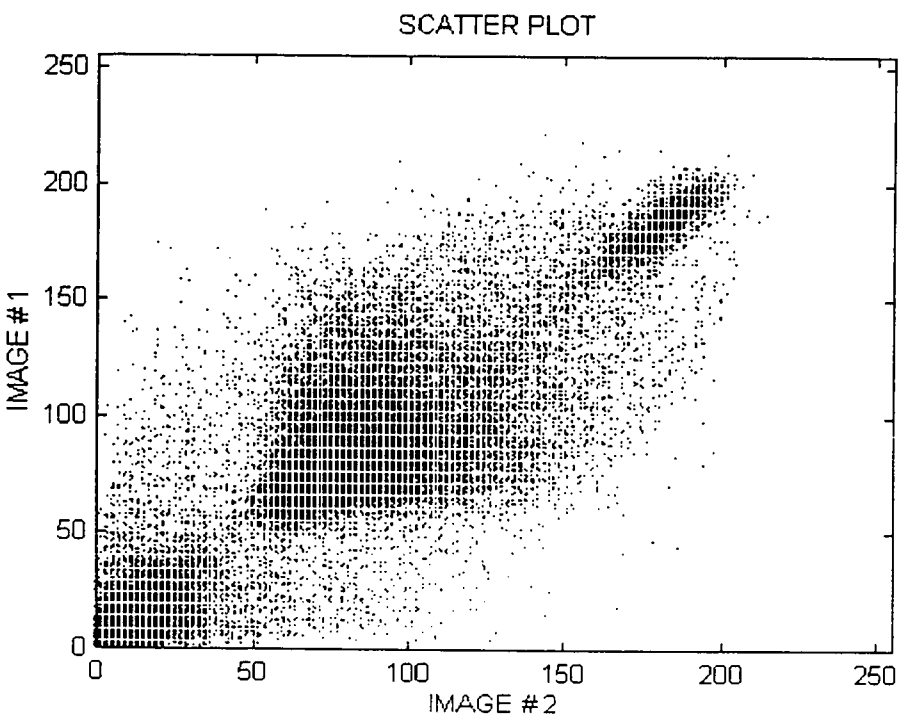


Figure 5-25 : Scatter plot of an MRI brain image with its rotated version (3^0).

Figure 5-26 shows the scatter plot of a brain image with a rotated version of the same image where the rotation is by 2^0.

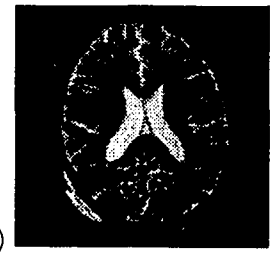
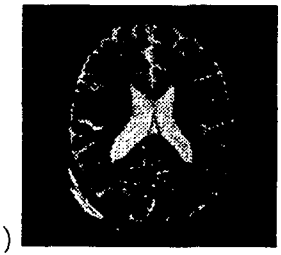
IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Image #1 rotated by 2^0) 

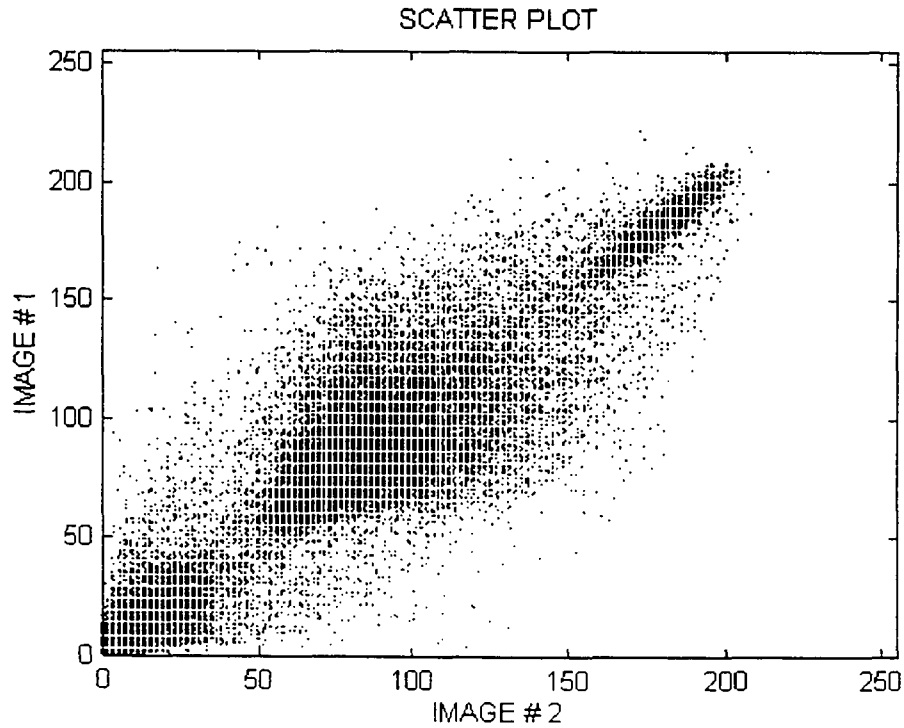


Figure 5-26 : Scatter plot of an MRI brain image with its rotated version (2^0).

Figure 5-27 shows the scatter plot of a brain image with a rotated version of the same image where the rotation is by 1^0.

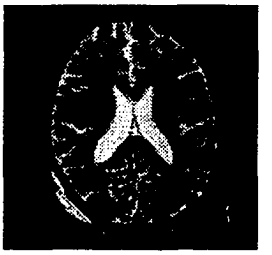
IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Image #1 rotated by 1^0) 

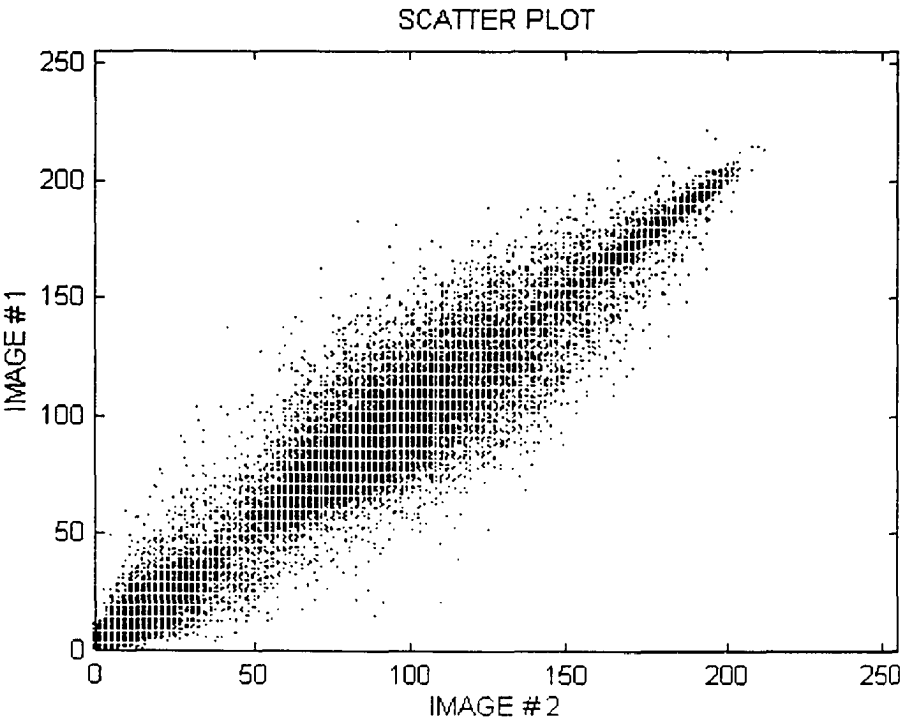


Figure 5-27 : Scatter plot of an MRI brain image with its rotated version (1^0).

When two images are perfectly registered, the cluster of points on the scatter plot is on a 45^0 line that passes through the origin. Figure 5-28 is a plot of two perfectly registered images.

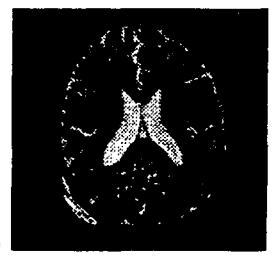
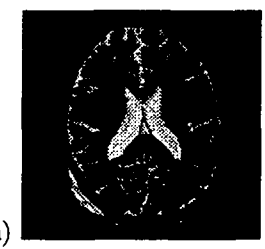
IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Same as IMAGE # 1. Ideal case of perfect registration) 

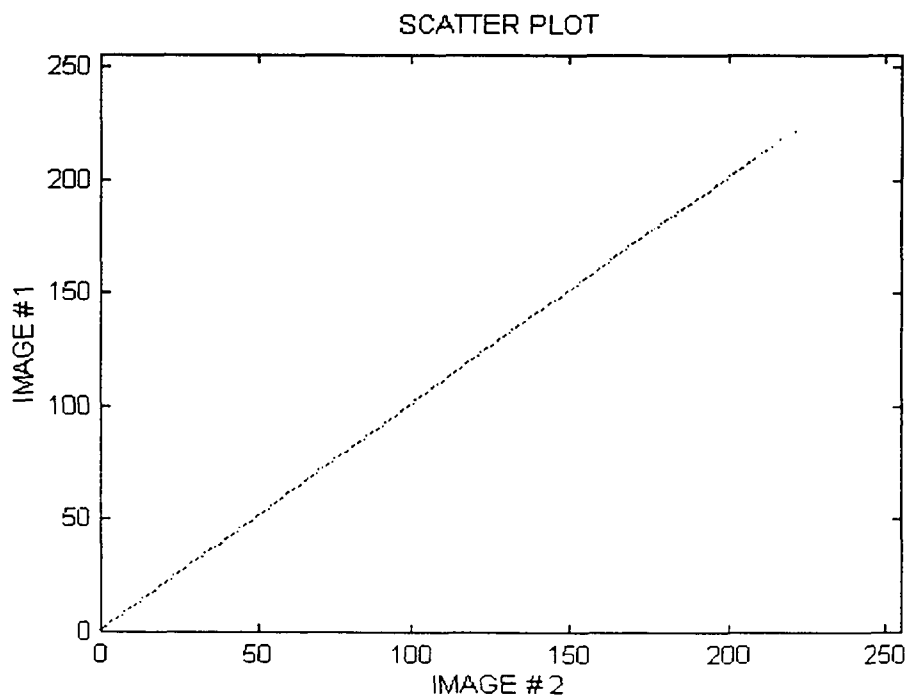


Figure 5-28 : Scatter plot of two perfectly registered MRI brain images

Figures 5-29 through 5-35 demonstrate the effect on the scatter plot due to spatial misalignment resulting from translation. Figure 5-29 shows a scatter plot of a MRI image of the brain with the same image translated by one pixel in the horizontal direction and one pixel in the vertical direction.

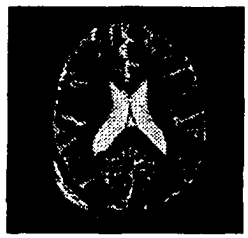
IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Translated version of IMAGE # 1) 

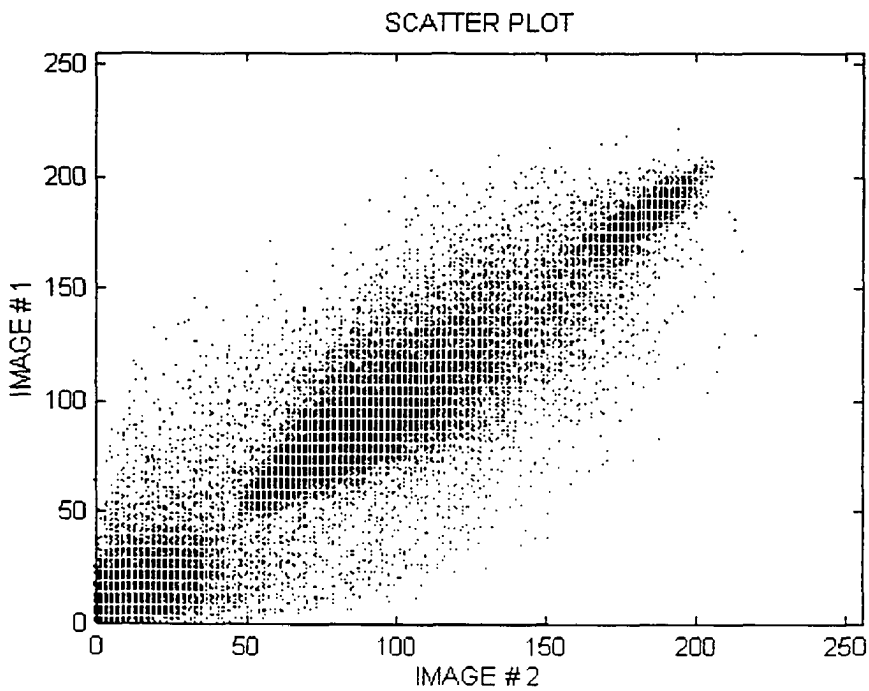


Figure 5-29 : Scatter plot of an MRI brain image with its translated version .
(Horizontal translation - 1 pixel and vertical translation - 1 pixel)

Figure 5-30 shows a scatter plot of a MRI image of the brain with the same image translated by two pixels in the horizontal direction and two pixels in the vertical direction.

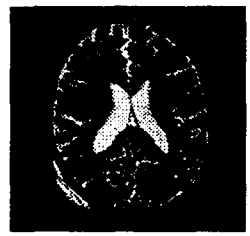
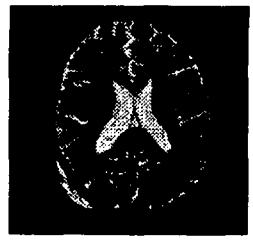
IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Translated version of IMAGE # 1) 

(Horizontal translation - 2 pixels and vertical translation - 2 pixels)

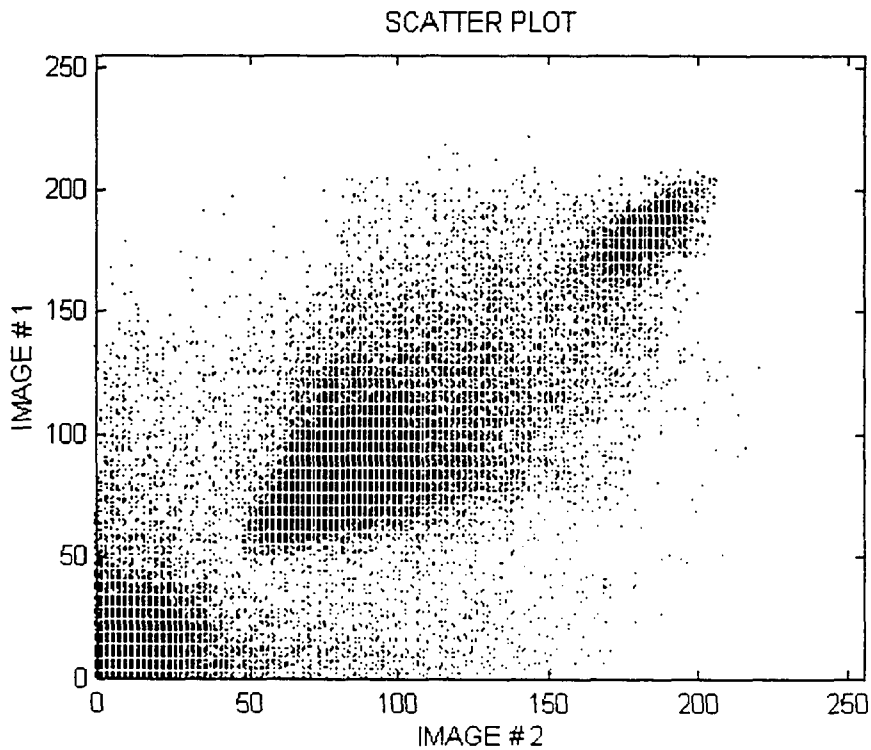
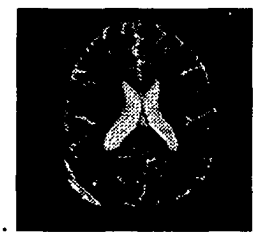


Figure 5-30 : Scatter plot of an MRI brain image with its translated version .
(Horizontal translation - 2 pixels and vertical translation - 2 pixels)

Figure 5-31 shows a scatter plot of a MRI image of the brain with the same image translated by three pixels in the horizontal direction and three pixels in the vertical direction.

IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Translated version of IMAGE # 1). 

(Horizontal translation - 3 pixels and vertical translation - 3 pixel)

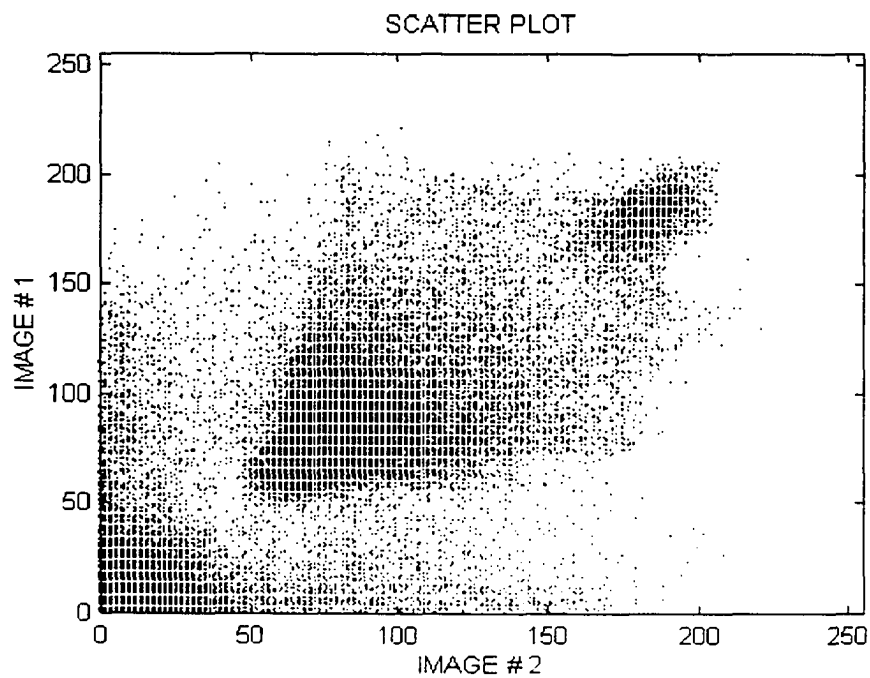
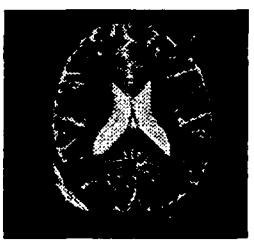


Figure 5-31 : Scatter plot of an MRI brain image with its translated version .
(Horizontal translation - 3 pixels and vertical translation - 3 pixels)

Figure 5-32 shows a scatter plot of a MRI image of the brain with the same image translated by four pixels in the horizontal direction and four pixels in the vertical direction.

IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Translated version of IMAGE # 1). 

(Horizontal translation - 4 pixels and vertical translation - 4 pixels)

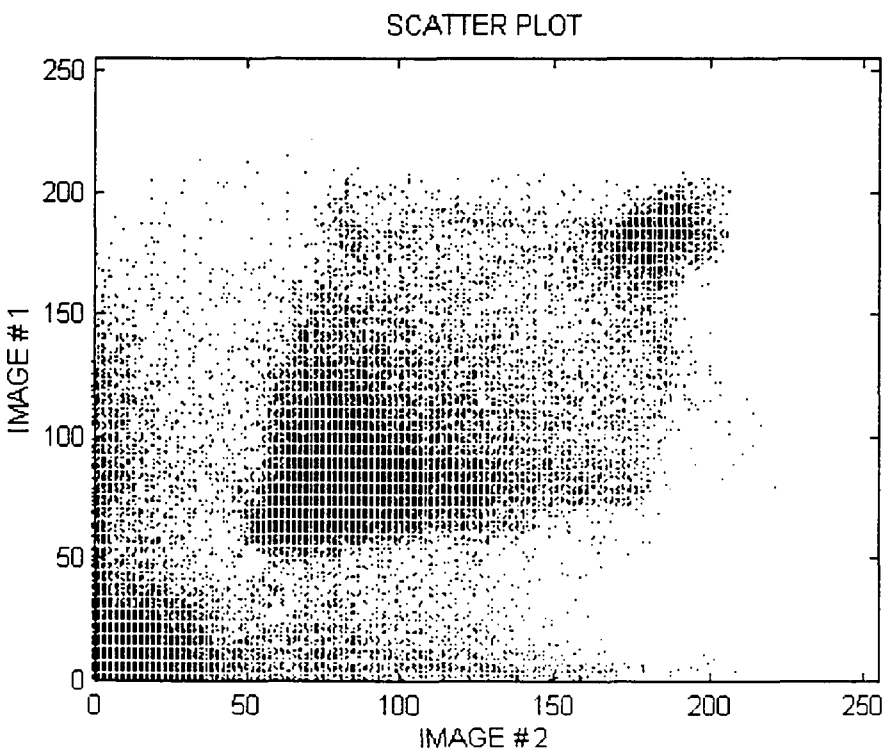
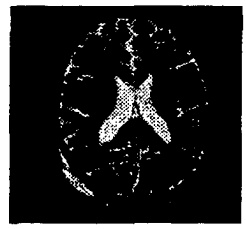


Figure 5-32 : Scatter plot of an MRI brain image with its translated version .
(Horizontal translation - 4 pixels and vertical translation - 4 pixel)

Figure 5-33 shows a scatter plot of a MRI image of the brain with the same image translated by five pixels in the horizontal direction and five pixels in the vertical direction.

IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Translated version of IMAGE # 1). 

(Horizontal translation - 5 pixels and vertical translation - 5 pixels)

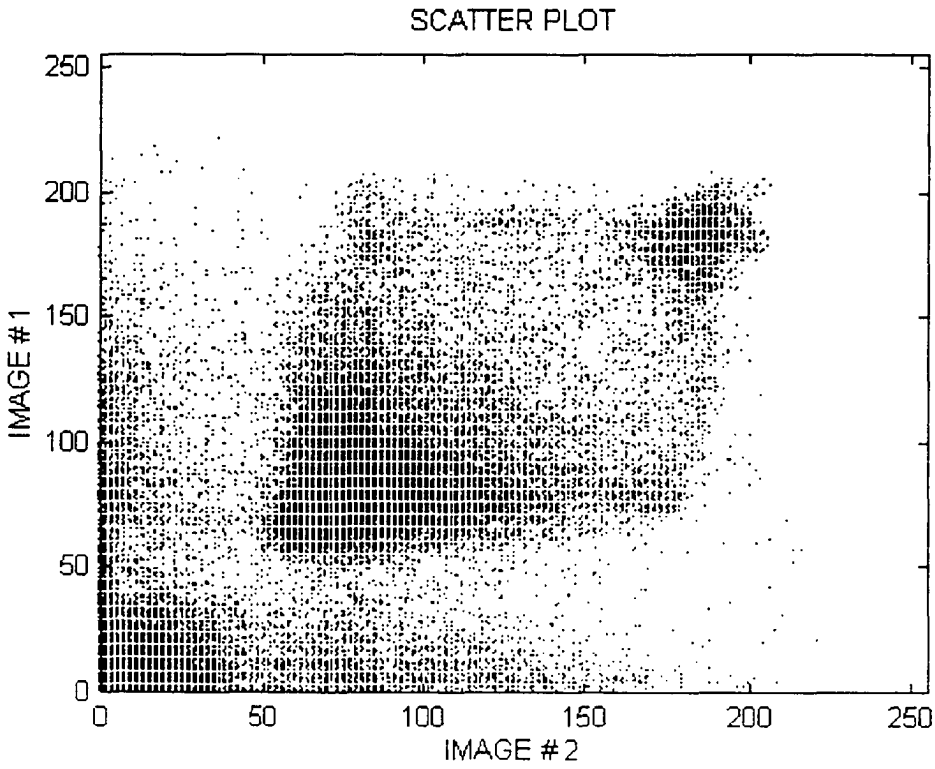
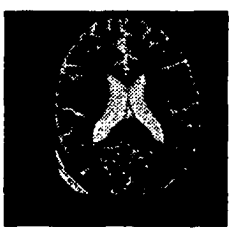


Figure 5-33 : Scatter plot of an MRI brain image with its translated version .
(Horizontal translation - 5 pixels and vertical translation - 5 pixels)

Figure 5-34 shows a scatter plot of a MRI image of the brain with the same image translated by ten pixels in the horizontal direction and ten pixels in the vertical direction.

IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Translated version of IMAGE # 1). 

(Horizontal translation - 10 pixels and vertical translation - 10 pixels)

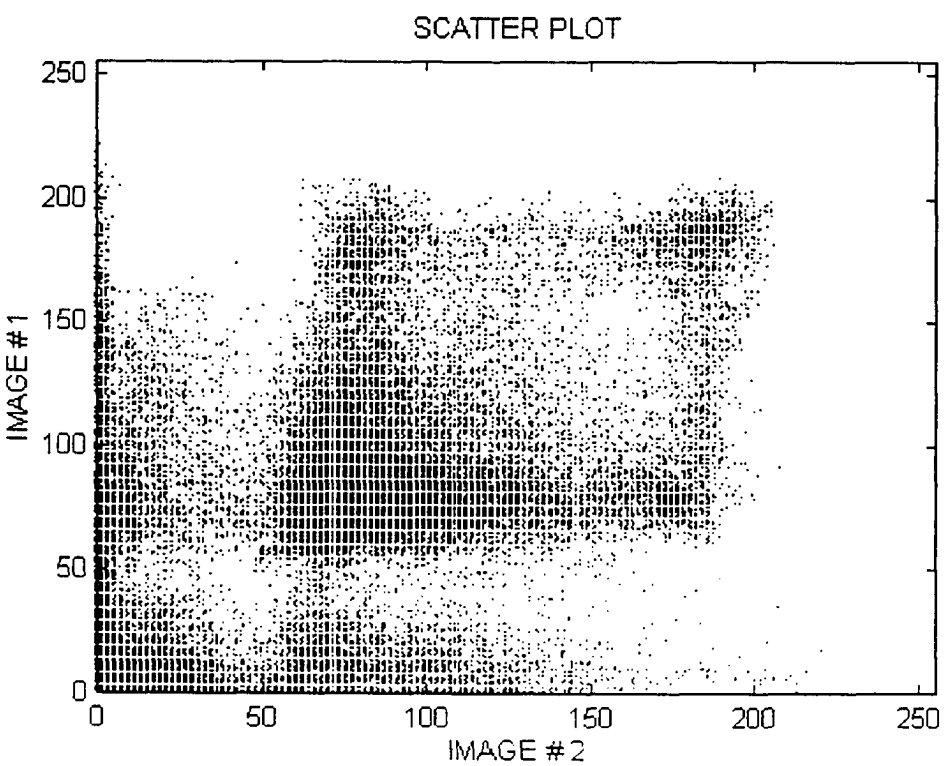


Figure 5-34 : Scatter plot of an MRI brain image with its translated version .
(Horizontal translation -10 pixels and vertical translation - 10 pixels)

Figure 5-35 shows a scatter plot of a MRI image of the brain with the same image translated by twenty pixels in the horizontal direction and twenty pixels in the vertical direction.

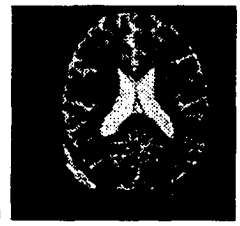
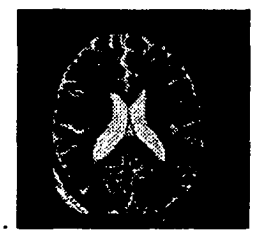
IMAGE # 1 (MRI of the brain) 

IMAGE # 2 (Translated version of IMAGE # 1). 

(Horizontal translation - 20 pixels and vertical translation - 20 pixels)

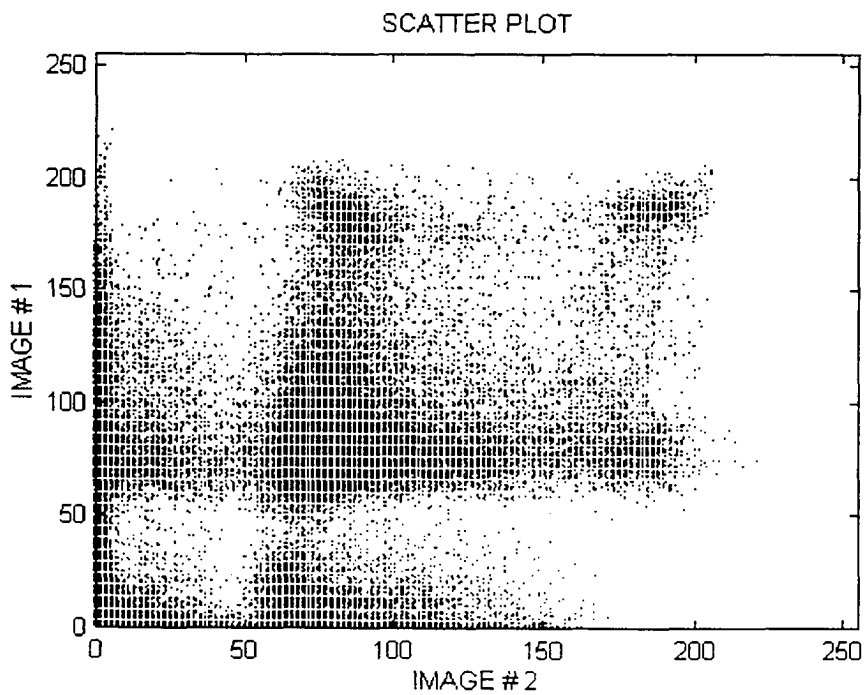


Figure 5-35 : Scatter plot of an MRI brain image with its translated version .
(Horizontal translation - 20 pixels and vertical translation - 20 pixels)

These scatter plots demonstrate that as the images are "more registered" the points on the scatter plot trend toward the 45^0 line through the origin.

Figure 5-40 shows a scatter plot for :

[1] Two rectangles shown in Figures 5-36 and 5-37 respectively, which overlap in the center of the image frame.

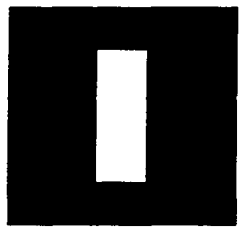


Figure 5-36 Image of a rectangle (IMAGE # 1).

Figure 5-37 : Image of a rotated but overlapping rectangle (IMAGE # 2).

[2] two squares shown in Figures 5-38 and 5-39 respectively, which overlap in the image frame.

Figure 5-38 Image of a rectangle (IMAGE # 1).

Figure 5-39 : Image of a shifted but overlapping square. (IMAGE # 2).

Note that gray level of zero is represented by black and gray level of 255 is represented by white. The scatter plot for the overlapping rectangles has the same distribution as the scatter plot for the overlapping squares.

An observation from the Figure 5-40 is that different images can have the same scatter plot. It also shows that the scatter plot does not give an indication in a quantitative sense

as to how well the images are registered. Scatter plots show trends and confirm how well objects are registered. However for some types of images scatter plots are suspect as a tool to determine how well images are registered. In the case of gray scale MRI images scatter plots do show trends that can be used to conclude how well images are registered.

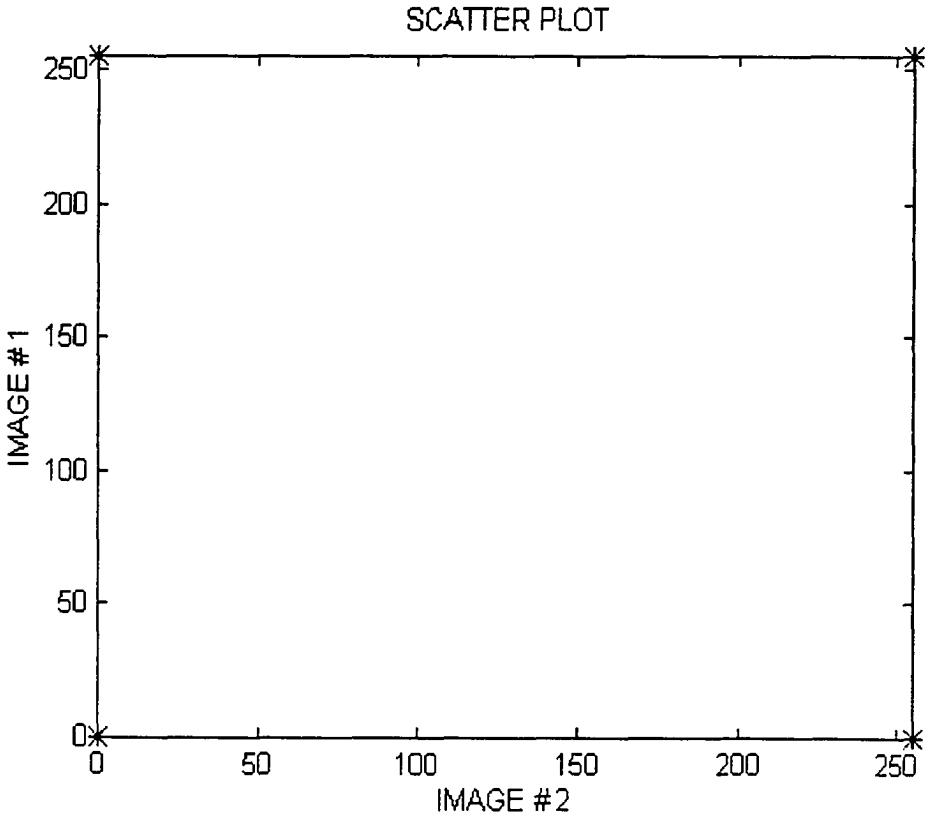


Figure 5-40 : Scatter plot
(1) for image in Figure 5-36 with image in Figure 5-37 and also
(2) for image in Figure 5-36 with image in Figure 5-37.

Figure 5-44 shows a scatter plot for :

[1] a rectangle shown in Figure 5-41 and a circle shown in Figure 5-42 respectively. These do not overlap.

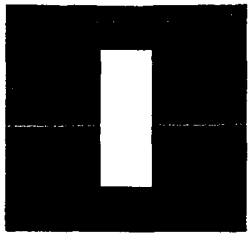


Figure 5-41 Image of a rectangle (IMAGE # 1).

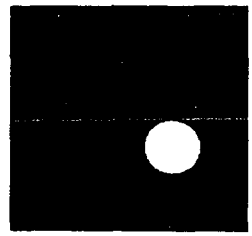


Figure 5-42 : Image of a non overlapping circle (IMAGE # 2).

[2] a square shown in Figure 5-43 and a circle shown in Figure 5-42 respectively. These do not overlap.

Figure 5-43 Image of a square (IMAGE # 1).

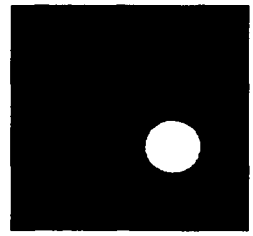


Figure 5-44 : Image of a non overlapping circle (IMAGE # 2).

Note that the scatter plot for the rectangle and the circle has the same scatter plot as the square and the circle. The scatter plot in Figure 5-45 clearly shows that scatter plot from dissimilar objects can have the same distribution.

Based on these scatter plots, we offer the following observations:

1. When images are registered the scatter plot is a 45^0 line through the origin.
2. When the registration error is within a few degrees (rotation) or a few pixels (translation), the scatter plot shows a significant clustering around a 45^0 line through the origin.
3. For similar images, these scatter plots cluster about a 45^0 line through the origin.

4. For dissimilar images, these scatter plots have no clustering about the 45^0 line through the origin.
5. A scatter plot cannot be used as a basis to predict the similarity of images. Dissimilar images can have the same scatter plot.
6. A scatter plot can provide a visual qualitative measure of how well images are registered, in most situations.

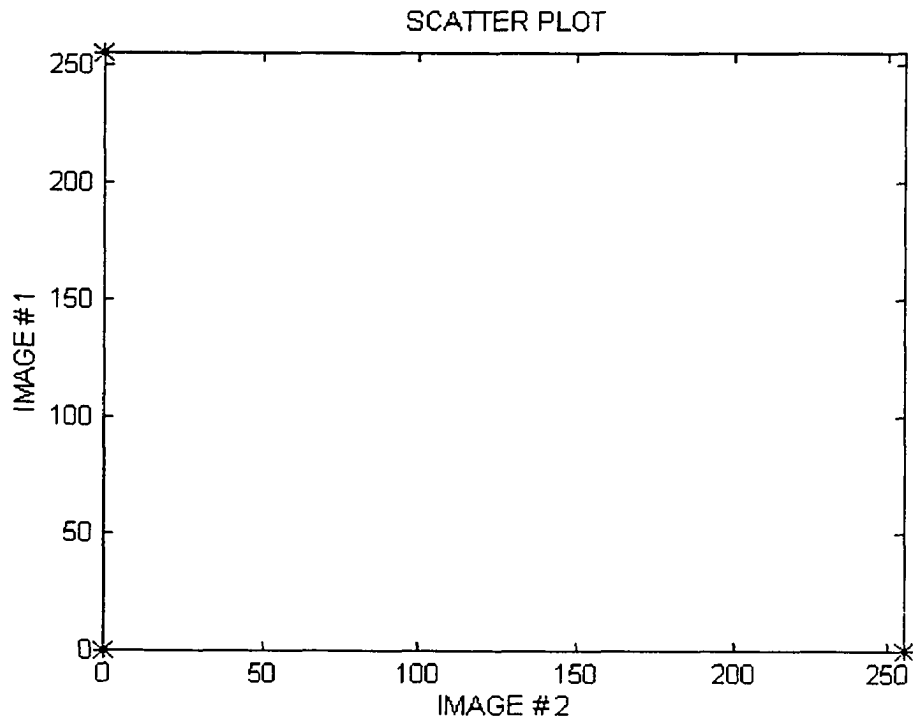


Figure 5-40 : Scatter plot
(1) for image in Figure 5-41 with image in Figure 5-42 and also
(2) for image in Figure 5-42 with image in Figure 5-43.

For certain images scatter plots are suspect as a tool. However for MRI images of the brain scatter plots do show strong trends as has been shown and explained. Additionally it provides a visual and qualitative measure.

5.4.4 Using a quantitative measure to validate

To validate a registration, perhaps the best measure is a quantitative measure. Consider an image $f_1(x,y)$. If this image $f_1(x,y)$ is translated and rotated by predetermined amount: is the registration technique able to determine the rotation and translation accurately? The technique used in this dissertation does.

If we now introduce artifacts inside the translated and rotated image (to simulate growth of cancer lesions in brain images); does the registration technique predict the rotation and translation correctly? The technique used in this dissertation does, with accuracy.

The correlation coefficient also provides a quantitative measure to estimate how well two images are registered. It suffers from some of the same deficiencies as the scatter plot. But since it is a quantitative measure it can be incorporated as part of an algorithm. This dissertation uses the correlation coefficient to determine how well images are registered. The next chapter is devoted to testing the registration method. Several cases have been considered and presented.

CHAPTER 6

EXPERIMENTAL RESULTS

6.1 Introduction

This chapter is devoted to presenting the results of the method of registration. Also presented are detailed steps taken to prepare the images for registration and the entire registration process as well as conclusions. It is important to point out that some of the figures in this chapter have been printed in reverse video format. In the normal image, the background is black and this corresponds to gray level 0 while white corresponds to the gray level of 255. In reverse video format, the gray level of 0 is painted as a white and black is changed to a gray value of 255. All other gray levels take on intermediary values. The reverse video format is obtained from the normal image by subtracting its gray levels from 255 and performing the display of the resultant image.

6.2 Validation

As previously mentioned, a demonstration that our method works is if we attempt to register images with known translations and rotations, and our procedure predicts the rotation and translation parameters correctly. Previously we showed that as images become more registered, the scatter plot of the images being registered converges towards the 45^0 line. This convergence provides a qualitative measure. For the same image, after registration with itself, the correlation coefficient is 1. Since we are considering similar images like MRI or CT images of the brain, aerial surveillance photos of the same region, the correlation coefficient between images being registered is non-negative (> 0) even between misaligned images and tends to increase to 1 as the images are better spatially

aligned. As the correlation coefficient r increases, r^2 increases. However this is not true in the general. The correlation coefficient between images can be negative and in these cases as r increases from -1 to 0, r^2 will actually decrease. For the class of images we considered for registration the correlation coefficient is non-negative and so r^2 increases when r does. In this research, the correlation coefficient is used as a measure to predict how well images are registered in place of the scatter plot. The correlation coefficient, r, between two images is defined as :

$$r = \frac{\sum_{x=0}^{x=IMAGESIZE-1} \sum_{y=0}^{y=IMAGESIZE-1} \left[f_1(x,y) - \overline{f_1(x,y)}\right]\left[f_2(x,y) - \overline{f_2(x,y)}\right]}{\left\{\sum_{x=0}^{x=IMAGESIZE-1} \sum_{y=0}^{y=IMAGESIZE-1} \left[f_1(x,y) - \overline{f_1(x,y)}\right]^2 \sum_{x=0}^{x=IMAGESIZE-1} \sum_{y=0}^{y=IMAGESIZE-1} \left[f_2(x,y) - \overline{f_2(x,y)}\right]^2\right\}^{1/2}} \quad \text{Eq. 6-1}$$

where

- $f_1(x,y)$ and $f_2(x,y)$ are the gray levels at each pixel location (x,y) in the first and second image being registered respectively, and
- $\overline{f_1(x,y)}$ and $\overline{f_2(x,y)}$ are the average gray levels in the first image and second image respectively, averaged over all image pixels in the image frame.

This measure is also referred to as Pearson's r. This term can have both negative and positive values. This is not desirable from the point of view of interpreting the value. The magnitude of r provides the relevant information in our case, indicating how well the images are registered. From a programming standpoint, r^2 is used as the measure to indicate how well two images are correlated. If r^2 is zero, the interpretation is that the images are not correlated and if a value of r^2 is one, the interpretation is complete correlation. Since the correlation coefficient is a quantitative measure, it is used as the

mechanism to stop iterations of our algorithm to register images. When r^2 exceeds a pre-assigned threshold value, we stop our algorithm and consider the images to be registered. Please note that when we register images we are registering objects like brain images from the same patients, etc. Registration is done for similar images. For example, it does not make clinical sense to register the image of a chest with a brain image. For this study, we define similar images from the perspective of the medical image registration application domain. We consider brain images from the same patient as similar images. However, it is recognized that similar images may constitute a broader and more inclusive definition to satisfy other application domains.

6.3 Results

In this section the steps in our registration are presented in pseudo code form to provide a programming environment as well as become more specific at each step. To establish that this procedure works and produces the necessary results within acceptable levels of tolerance, results are presented for the following two categories:

1. Simple pictures.
2. MRI Brain Images.

We consider translation and rotation as well as the introduction of artifacts to show that in each of the cases this procedure does register.

6.3.1 Pseudo Code

Procedure Registration(image1, image2)

```
Begin
    c_image1 = Center_Image_in_Image_frame (image1);
    c_image2 = Center_Image_in_Image_frame (image2);
```

```
corr = correlation_coefficient (c_image1, c_image2);
Rotation_angle = 0;
if (corr < Threshold1) {
   principal_axis1 = Find_PCS_after_Modmax_and_Hull(c_image1);
   temp_image = c_image2;
   do {
      principal_axis2 = Find_PCS_after_Modmax_and_Hull(temp_image);
      Angle = Compute_angle(principal_axis1, principal_axis2)
      Rotation_angle = Rotation_angle + angle;
      temp_image = Rotate_second_image(c_image2, Rotation_angle)
      corr = correlation_coefficient (c_image1, temp_image);
   } while (corr < Threshold1)
}
Do Fine_tuning(c_image1, temp_image, corr);
End
```

Procedure Find_PCS_after_Modmax_and_Hull (image);

```
Begin
```

1. Modmax_Image = Wavelet_Modmax_program(image);
2. Hull = Convex_Hull(Modmax_image);
3. Principal_axis = Principal_Components (Hull);

```
End
```

Procedure Center_Image_in_Image_frame (image)

```
Begin
```

```
    Do twice the following 4 steps {
        1. Modmax_Image = Wavelet_Modmax_program(image);
        2. Hull = Convex_Hull(Modmax_image);
        3. Centeroid_Coordinates = Find_centeroid(Hull);
        4. image = Translate_image_to_Frame_Center (Centroid_Coordinates,
                                                    IMAGESIZE, image);
    }
End
```

Procedure Fine_Tuning (image1, image2, corr)

```
Begin
    For dir = (curr, north, northeast, east, southeast,
               south, southwest, west, northwest) {
        LABEL1 : T_IMG = Translate_image(image2, dir, 1 pixel);
                 new_corr = correlation_coefficient (image1, T_IMG);
                 if (new_corr > corr) {
                     image2 = T_IMG;
                     corr = new_corr;
                     while {
                         R_IMG = Rotate(image2, counter_clockwise, 1 degree);
                         new_corr = correlation_coefficient (image1, R_IMG);
                         if (new_corr > corr) {
                             image2 = R_IMG;
                             corr = new_corr;
```

```
                    }
                    else {
                        R_IMG = Rotate(image2, clockwise, 1 degree);
                        new_corr = correlation_coefficient (image1,
                        R_IMG);
                        if (new_corr > corr) {
                        image2 = R_IMG;
                        corr = new_corr;
                    }
                }
                do (corr > new_corr);
                GOTO LABEL1;
            }
        }
    END
```

6.3.2 Simple pictures

Several cases are considered which are presented case by case.

6.3.2.1 Case 1

Figures 6.1 and 6.2 are images of two rectangles. Figure 6.2 is a translation of the same rectangle shown in Figure 6.1.

The horizontal edges, vertical edges and the Wavelet Modulus Maxima Images are generated by the wavelet analysis program on the object in the images shown in Figures 6-1 and 6-2. A hull program uses the points that constitute the Wavelet Modulus Maxima Image as input and generates the convex hull for each of the Modulus Maxima Images. An

image centering algorithm takes the hull points, computes the centroid and moves each of the images to the center of the image frame. Table 6-1 shows the relationship of subsequent Figures 6-3 through 6-12 to the images shown in Figures 6-1 and 6-2. Subsequent Figures 6-3 through 6-12 present computed results for Figures 6-1 and 6-2.

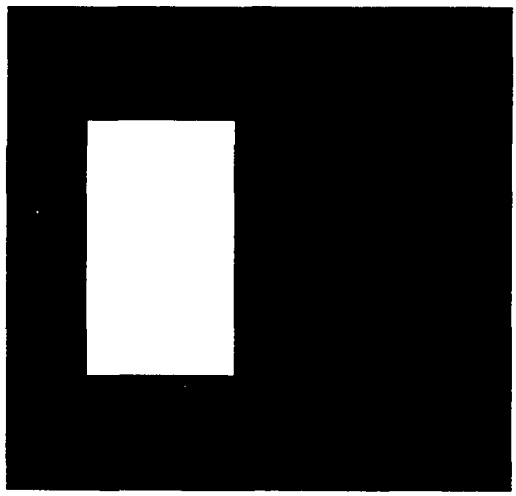


Figure 6-1 : Rectangle

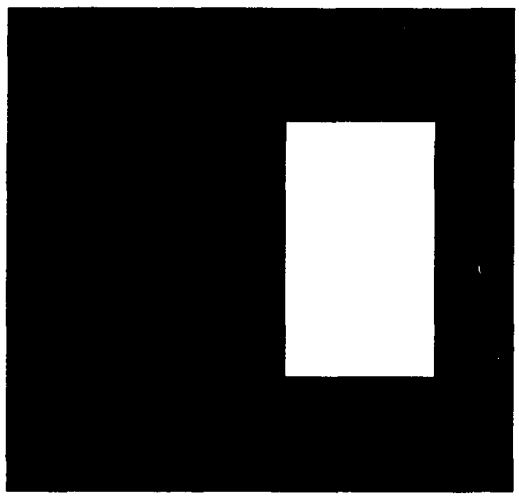


Figure 6-2 : Translated Rectangle

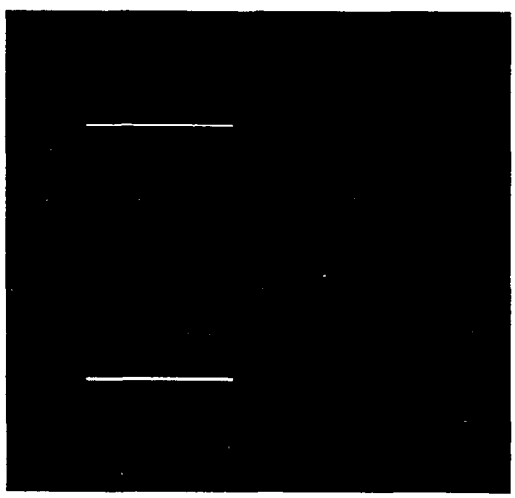


Figure 6-3 : Horizontal Edge (Image # 1)

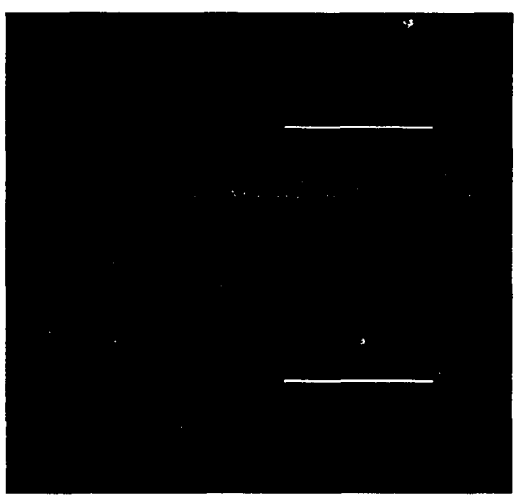


Figure 6-4 : Horizontal Edge (Image # 2)

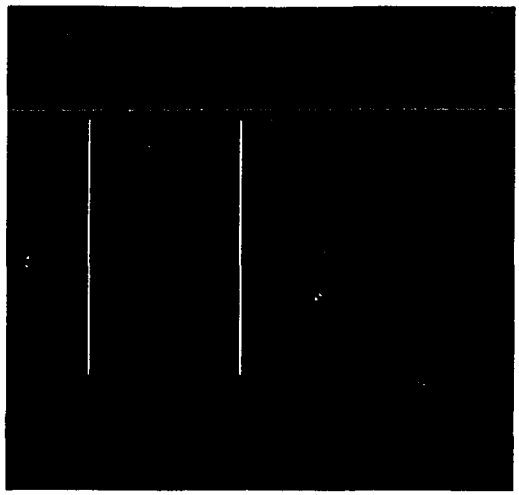


Figure 6-5 : Vertical Edge (Image # 1)

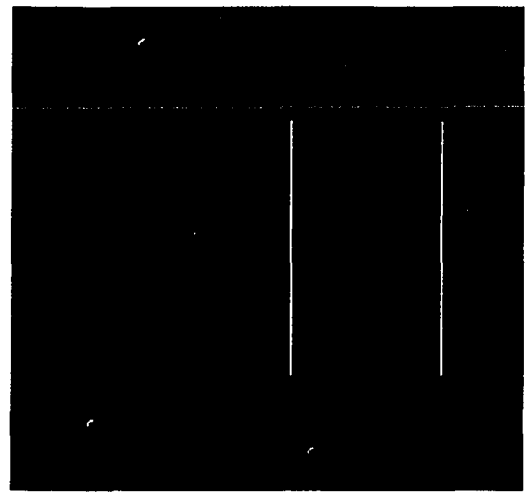


Figure 6-6 : Vertical Edge (Image # 2)

Table 6-1 : This table shows the relationship of subsequent Figures 6-3 through 6-12 to the images shown in Figures 6-1 and 6-2.

Image type	Image #1 shown in Figure 6-1	Image #2 shown in Figure 6-2
Vertical Edge	Figure 6-3	Figure 6-4
Horizontal Edge	Figure 6-5	Figure 6-6
Modulus Maxima	Figure 6-7	Figure 6-8
Hull	Figure 6-9	Figure 6-10
Registered Image	Figure 6-11	Figure 6-12

The correlation coefficient (r) and the correlation measure (r^2) for the images in Figures 6-11 and 6-12 are presented in Table 6-2.

Table 6-2 : The correlation coefficient (r) and the correlation measure (r^2) for the images in Figures 6-11 and 6-12.

Correlation coefficient (r)	1.000000
Correlation measure r-square	1.000000

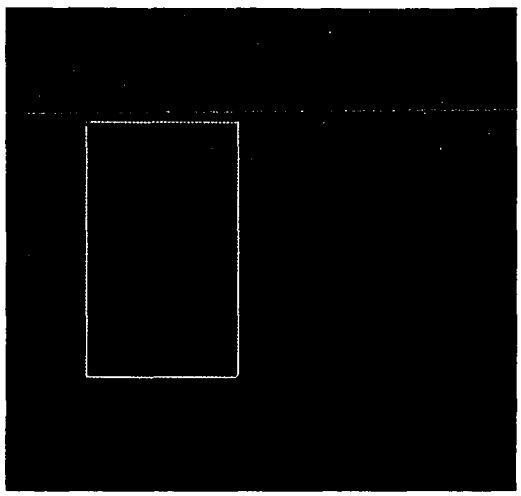


Figure 6-7 : Modulus Maxima Image (Image # 1)

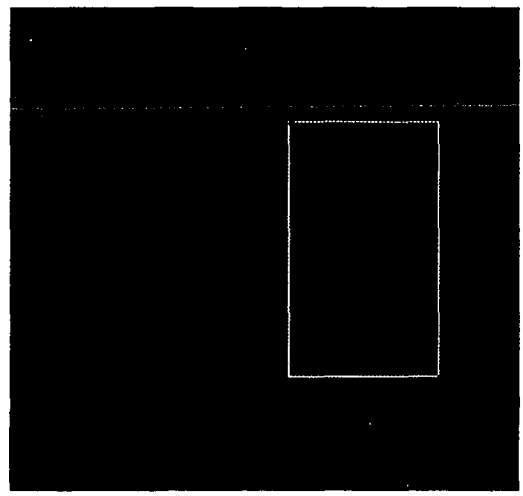


Figure 6-8 : Modulus Maxima Image (Image # 2)

Figure 6-9 Hull : (Image #1)

Figure 6-10 : Hull (Image # 2)

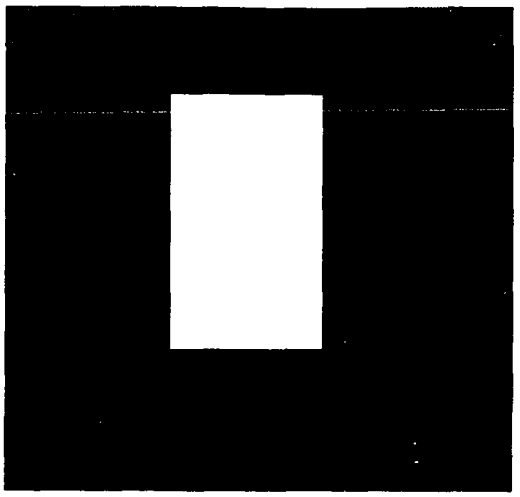


Figure 6-11 : Image # 1 after registration

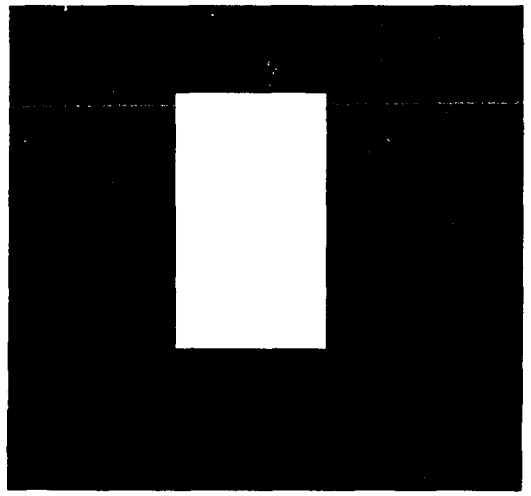


Figure 6-12 : Image # 2 after registration

6.3.2.2 Case 2

Figures 6-13 and 6-14 are images of the same rectangle. The rectangle in Figure 6-14 is created by a rotation of the rectangle in Figure 6-13 of ninety degrees.

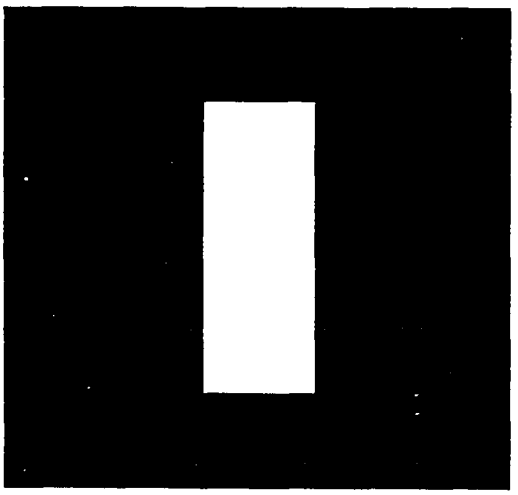


Figure 6-13 : Image # 1 .

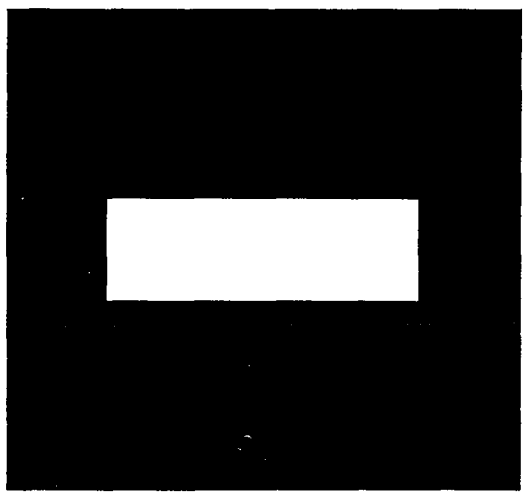


Figure 6-14 : Image # 2 - rotated version of Image # 1.

Figure 6-15 shows the difference image for the purpose of illustration.

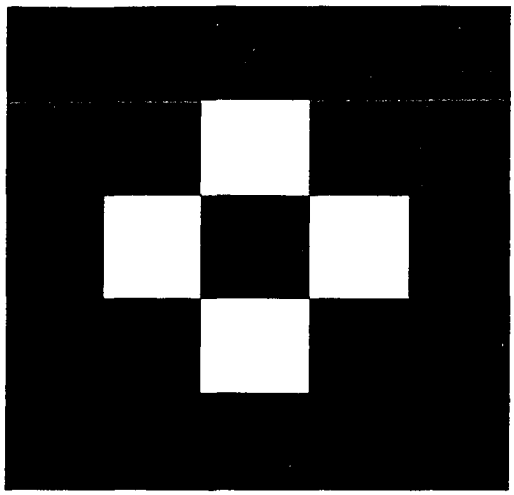


Figure 6-15 : Difference image between image # 1 and # 2.

A Wavelet Modulus Maxima image is obtained for each of the images in Figures 6-13 and 6-14. These points on the Modulus Maxima image serve as input to a program that computes the convex hull.. The convex hull program provides us the points that are shown in Figures 6-16 (hull points for image in figure 6-13) and 6-17 (hull points for image in figure 6-14). The principal component analysis program uses the points on the convex hull as input and computes the angle the principal axis makes with the horizontal direction. For the points displayed in Figures 6-16 and 6-17 this angle was computed to be 92.46^0 and -2.18^0 respectively. Based on this information the angle of rotation is computed to be 94.64^0. The image in Figure 6-18 is obtained by rotating the image in figure 6-14 by 94.64^0 in the counterclockwise direction. Figure 6-19 shows the difference image obtained by subtracting the image in Figure 6-18 from the image in Figure 6-13. The expectation was a resultant angle of rotation of 90^0, but this did not happen. An analysis of the hull, shows that each of the hulls have five points, not four as may be visually apparent.

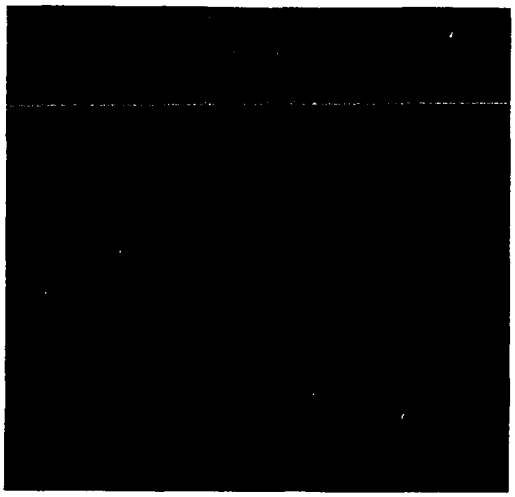


Figure 6-16 : Hull (Image # 1)

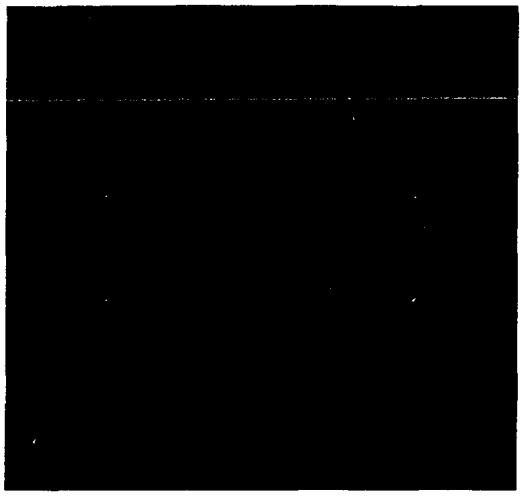


Figure 6-17 : Hull (Image # 2)

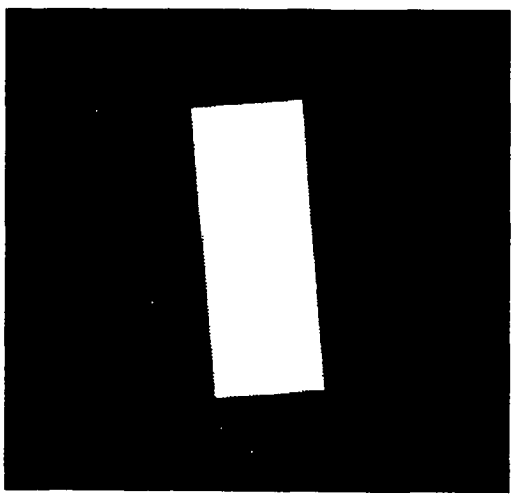


Figure 6-18 : Image # 2 rotated by 94.63^{0}.

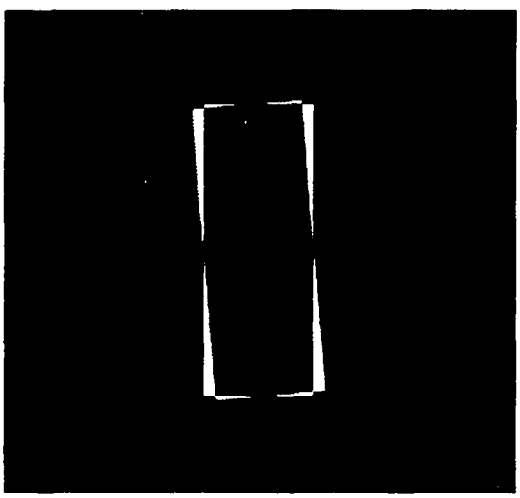


Figure 6-19 : Difference Image between Image # 1 (Figure 6-13) and image in Figure 6-18.

These five points include the bottom left corner of the hull which actually consists of two coordinate points or two pixels. This results in a principal axis, determined by the principal component analysis, to be skewed to 92.46^{0} and -2.18^{0}. This skewing produces an angle of rotation of 94.64^{0} instead of the expected (correct) rotation of 90^{0}. Further analysis of this problem reveals that when we do wavelet analysis, we essentially perform a

convolution. A convolution actually extends the signal and in a sense "walks" the edges. This effect is not noticeable in Figures 6-16 and 6-17 with the eye. In filter theory, such a distortion means an ideal delay in a Linear Time Invariant (LTI) system. In wavelet analysis, there are two phases, an analysis phase and a synthesis phase. The synthesis phase has a built in delay mechanism to deal with this delay. However, if we consider only the analysis phase, the edge gets shifted. Moreover, an analysis of many edge detection algorithms [Weeks97][Dougherty87] that involve the use of a mask, e.g. the Sobel edge detector, Prewitt Edge detector, Robert's Edge detector have their edges move. This is analogous to what a convolution produces. With wavelets, this directs attention to two issues:

1. The shorter the high pass filter , the fewer the number of pixels the edge moves. The movement as far the modulus maxima image is concerned also depends on the values of the filter coefficients.
2. As a wavelet analysis is done over scale, the convolution causes the edge to move at each stage. This movement of the edge can be expressed in terms of filter lengths and stage of the iteration. This edge movement expression is as follows:

   Let

   n_1 denote the length of the lowpass filter.

   n_2 denote the length of the high pass filter.

   stg denote the stage of the iteration in the wavelet analysis process.

- Then at each stage of the wavelet analysis, the lowpass image blurs to envelop $(n_1 - 1)$ extra pixels. The resultant edge movement for low pass images is $(n_1 - 1) * stg$.

- The high pass edges are detected at most $(n_2 - 1)$ pixels away from the actual edges in the first stage. In subsequent iterations, the high pass edges move out by a factor of at most $(n_1 - 1) * stg + (n_2 - 1)$. We say "at most", because this movement also depends upon the filter coefficient values. This high pass edge contributes to the detection of the modulus maxima edges. Hence, the modulus maxima edges also move by at most $(n_1 - 1) * stg + (n_2 - 1)$ pixels. An important consideration that needs to be accounted for, is the size of the pixels.
- The direction of the edge movement is in the direction in which the input signal is processed.

Figure 6-20 show the Modulus Maxima Image for the image in Figure 6-18 and Figure 6-21 shows the convex hull for the image in Figure 6-20.

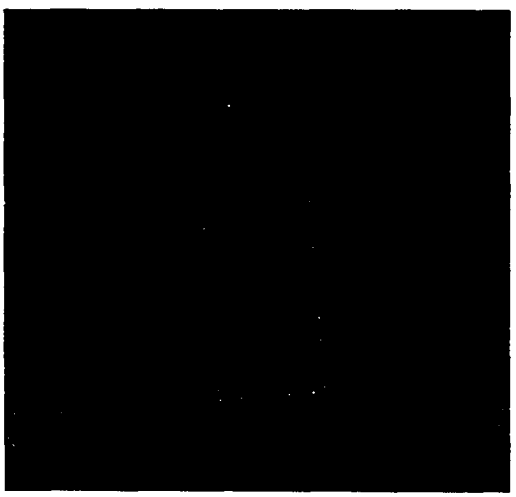


Figure 6-20 : Modulus Maxima Image for image in Figure 6-18.

Figure 6-21 : Hull points from Modulus Maxima Image in Figure 6-20.

The rectangle in Figure 6-20 does not show a continuous line but a line that is a string of several small islands of dashes. The reason for this is the integer approximation of the pixel coordinates. The Wavelet program does find the points. If the resolution is much higher these small dashed lines seen in Figure 6-20 would appear more closer together

giving the perception of a continuous line. The correlation coefficient between the images in Figures 6-13 and 6-18 is 0.9341. After computing the principal components for the hull points shown in Figure 6-21 we find that the principal axis is oriented at 91.15^0 with the horizontal. Since the principal axis for the hull points shown in Figure 6-16 (for IMAGE # 1) makes and angle of 92.46^0 with the horizontal direction our program determines that the rotation angle should be an additional 1.31^0. Figure 6-14 is now rotated by 95.95^0 ($94.64^0 + 1.31^0$) in the counter clockwise direction and the rotated image is shown in Figure 6-22. Figure 6-19 shows the difference image obtained by subtracting the image in Figure 6-18 from the image in Figure 6-13.

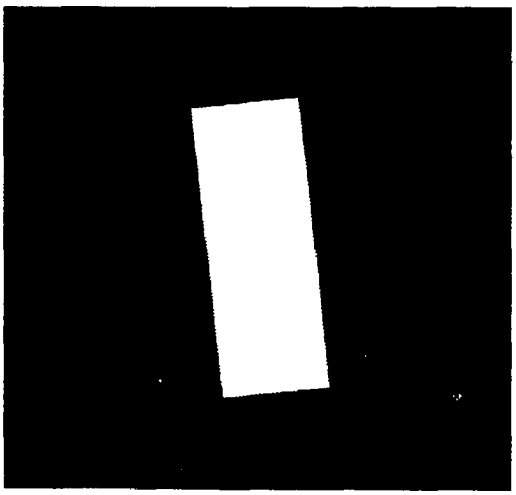


Figure 6-22 : Image in Figure 6-18 rotated by 95.95^0.

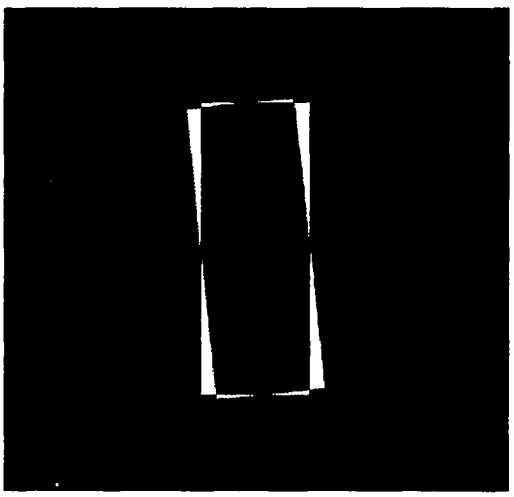


Figure 6-23 : Difference image between Image in Figures 6-13 and 6-22.

The correlation coefficient between the Figures 6-13 and 6-22 computes to 0.915. As the value is less than the correlation coefficient between the images in Figures 6-13 and 6-18 is above the threshold and higher than the value obtained from the current iteration the registration program uses the image in Figure 6-18 and performs a fine tuning. In the fine tuning stage at each iteration the figure is rotated by 1^0 in the clockwise or counter

clockwise direction. If the correlation coefficient is improves the step is continued in the same direction. However the rotation is retracted if the correlation coefficient decreases from its previous value. The fine tuning algorithm also translates the image by 1 pixel in each of the possible directions to improve the correlation coefficient, retracting a step if the correlation coefficient drops. Figures 6-24, 6-26, 6-28, 6-30 and 6-32 are obtained by successively rotating by one additional degree the image in Figure 6-18 in the clockwise direction. This is achieved by actually rotating the image in Figure 6-14 by 93.64^0, 92.64^0, 92.64^0 and so on in the counter clockwise direction. The reason for doing this is because rotation causes geometric distortion. This rotation is done by the fine tuning program since each rotation by 1^0 led to an increase in the correlation coefficient between the rotated image and the image in Figure 6-13.

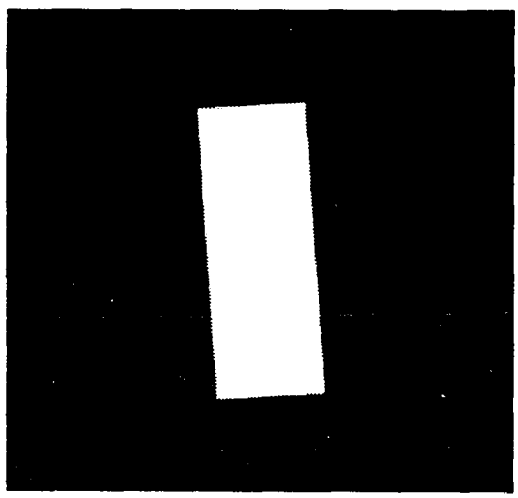


Figure 6-24 : Figure 6-14 rotated by 93.64^0 (counter clockwise rotation). (Fine Tuning Program iteration).

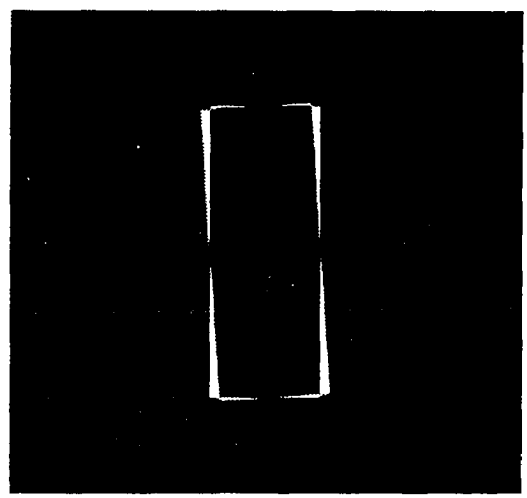


Figure 6-25 : Difference image between image in Figures 6-13 and 6-24.

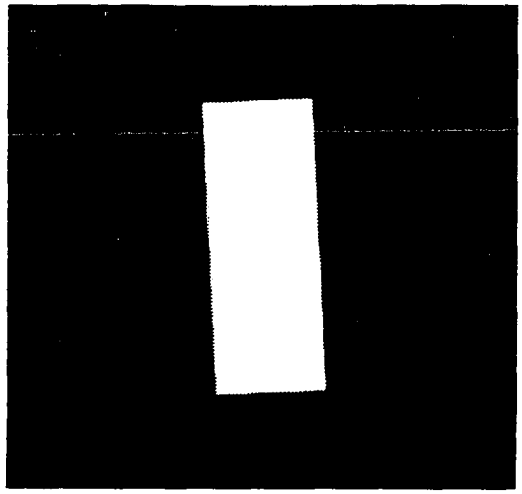


Figure 6-26 : Figure 6-14 rotated by 92.64^{0} (counter clockwise rotation). (Fine Tuning Program iteration).

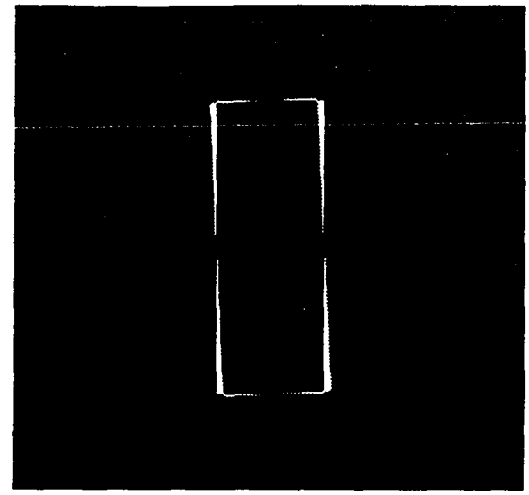


Figure 6-27 : Difference image between image in Figures 6-13 and 6-26.

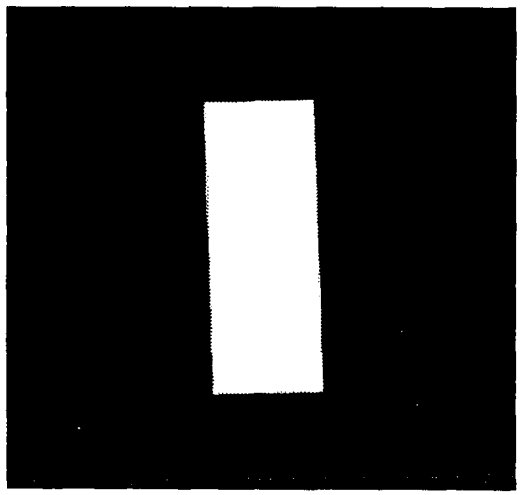


Figure 6-28 : Figure 6-14 rotated by 91.64^{0} (counter clockwise rotation). (Fine Tuning Program iteration).

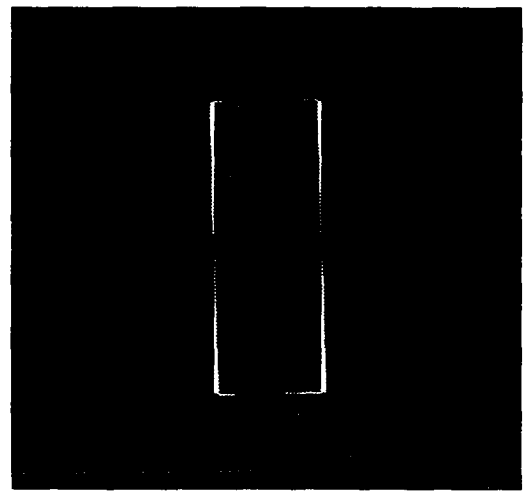


Figure 6-29 : Difference image between image in Figures 6-13 and 6-28.

As the difference image shows, and the value of the correlation coefficient (r) confirms, our algorithm does register the images in Figures 6-13 and 6-14 correctly. For this test we

used a fixed rotation angle of 90^0 and the algorithm determines the angle of rotation to be 89.64^0.

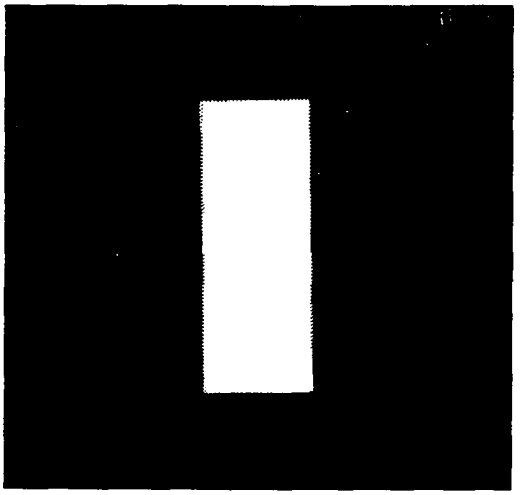


Figure 6-30 : Figure 6-14 rotated by 90.64^0 (counter clockwise rotation). (Fine Tuning Program iteration).

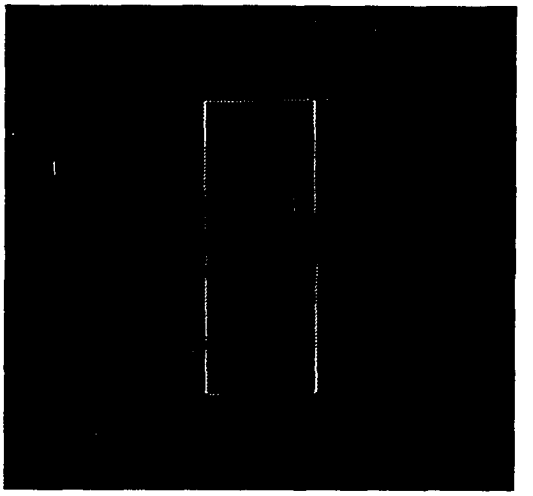


Figure 6-31 : Difference image between image in Figures 6-13 and 6-30.

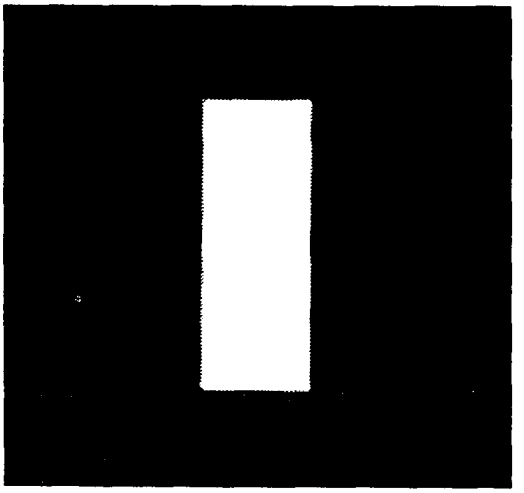


Figure 6-32 : Figure 6-14 rotated by (counter clockwise rotation). (Fine Tuning Program iteration).

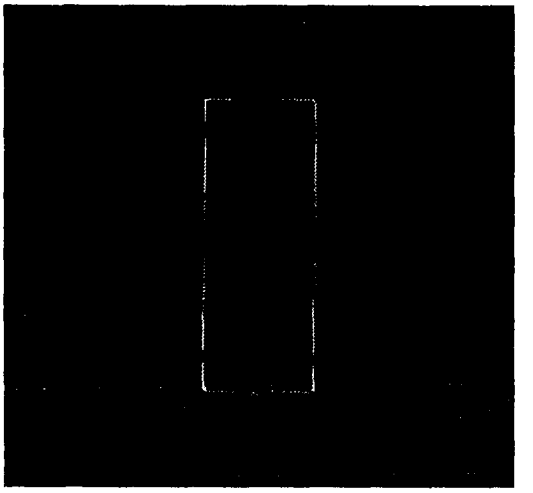


Figure 6-33 : Difference image between image in Figures 6-13 and 6-32.

Table 6-3 provides the correlation coefficient as well as the r^2 measure for each stage of the iteration.

Table 6-3 : Table showing the correlation coefficient and measure with each iteration.

Image and Iteration	**Correlation coefficient (r) with image # 1**	**Correlation measure (r^2) with image # 1**
Original Image # 2 (At 90^0 to image #1)	0.258369	0.066755
Image # 2 rotated by 94.64^0 [counter clockwise]	0.934102	0.872547
Image # 2 rotated by $95.95^0 (= 94.64^0 + 1.31^0)$ [counter clockwise]	0.915322	0.837815
Image # 2 rotated by $93.64^0 (= 94.64^0 - 1^0)$ [counter clockwise](Fine Tuning Program)	0.9541194	0.904770
Image # 2 rotated by $92.64^0 (= 94.64^0 - 1^0 - 1^0)$ [counter clockwise](Fine Tuning Program)	0.967081	0.935245
Image # 2 rotated by $91.64^0 (= 94.64^0 - 1^0 - 1^0 - 1^0)$ [counter clockwise](Fine Tuning Program)	0.981562	0.963464
Image # 2 rotated by $90.64^0 (= 94.64^0 - 1^0 - 1^0 - 1^0 - 1^0)$ [counter clockwise](Fine Tuning Program)	0.992640	0.985335
Image # 2 rotated by $89.64^0 (= 94.64^0 - 1^0 - 1^0 - 1^0 - 1^0 - 1^0)$ [counter clockwise](Fine Tuning Program)	0.994092	0.988219

6.3.2.3 Case 3

Figures 6-34 and 6-35 are images of the same rectangles, however this time, the rectangle in Figure 6-35 is a translation of the rectangle in Figure 6-34. Additionally some artifacts were intentionally added into the rectangle. In brain images for example, there may be growth or shrinkage in tumors and this should not affect the registration. In bone grafts, the bones may integrate and this may cause volume shrinkage. The registration should work despite these internal changes. This algorithm also provides good results in such cases. The shape and size of the internal changes does not affect the accuracy of registration. The registration technique has been made insensitive to the internal changes. Even when the two images are spatially aligned, the correlation coefficient may not reach high levels. The extent of internal change will also dictate the correlation coefficient. So instead of judging how well images are registered by a threshold value it is important to also keep track of the point where the correlation coefficient reaches a maxima value.

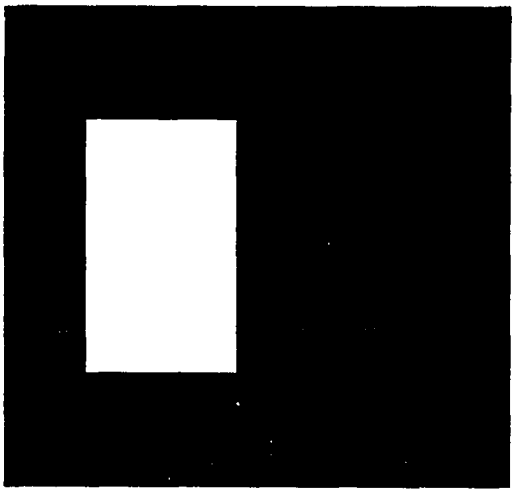


Figure 6-34 : Image #1 Rectangle.

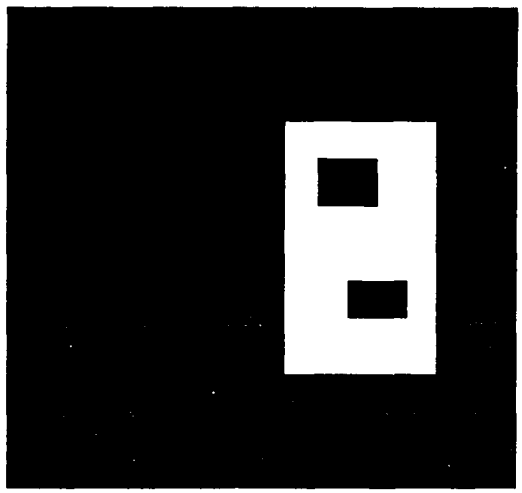


Figure 6-35 : Image # 2 .

Again the Modulas Maxima Image and the convex hull points are computed for each of the images shown in Figures 6-34 and 6-35.

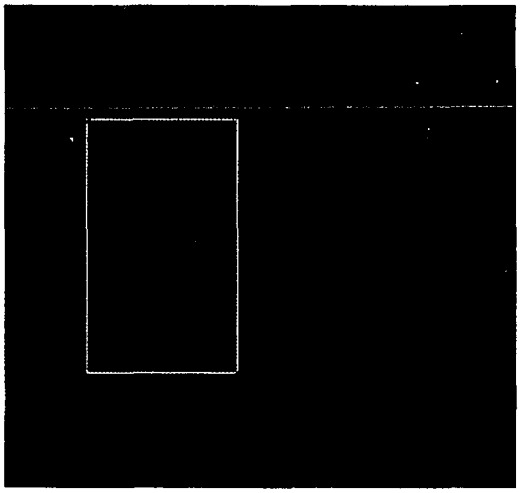


Figure 6-36 : Modulus Maxima Image for for Image # 1 (Figure 6-34).

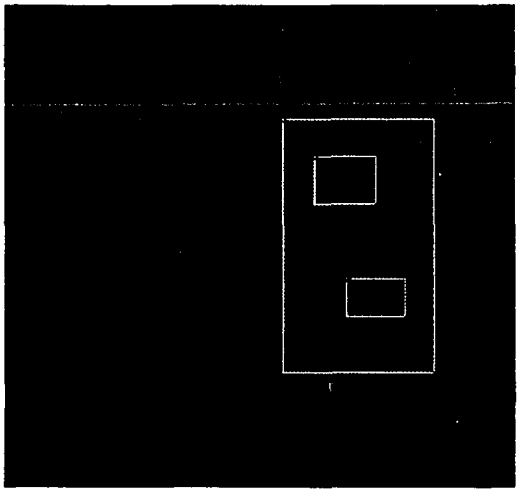


Figure 6-37 : Modulus Maxima Image for for Image # 2 (Figure 6-35).

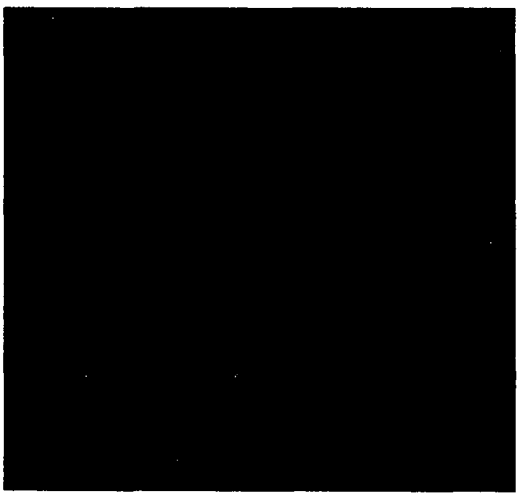


Figure 6-38 : Hull points for image in Figure 6-36.

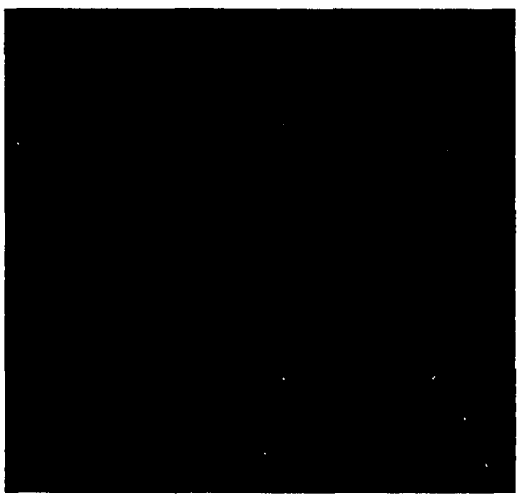


Figure 6-49 : Hull points for image in Figure 6-37.

Figures 6-36 and 6-37 are the Modulas Maxima Images and Figures 6-38 and 6-39 are the convex hull points for the images in Figures 6-34 and 6-35 respectively. Then the image frame centering program is run using these hull points shown in Figures 6-38 and 6-39, as input. This program finds the centroid of the hull points and computes the translation parameters to translate each of the images to the center of the image frame.

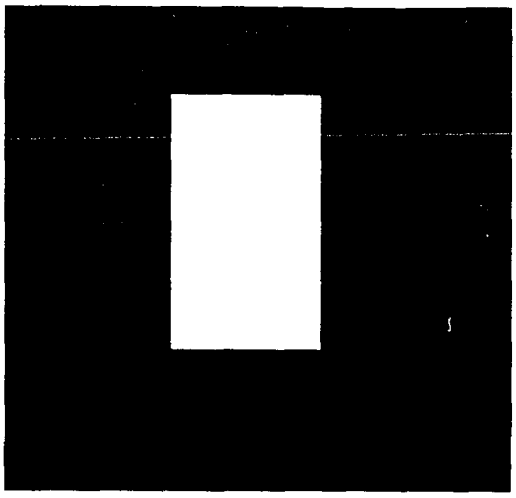


Figure 6-40 : Image # 1 shown in Figure 6-34 centered in image frame.

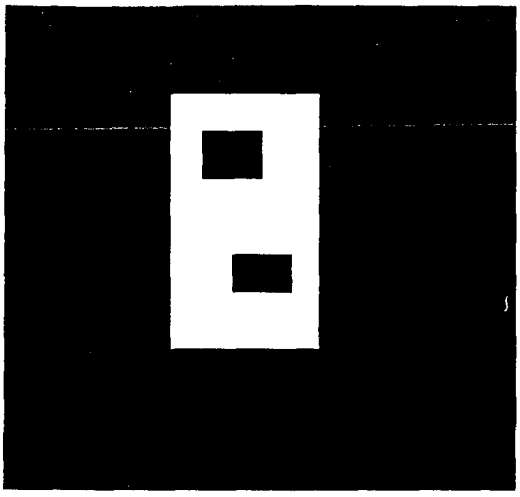


Figure 6-41 : Image # 2 shown in Figure 6-35 centered in image frame.

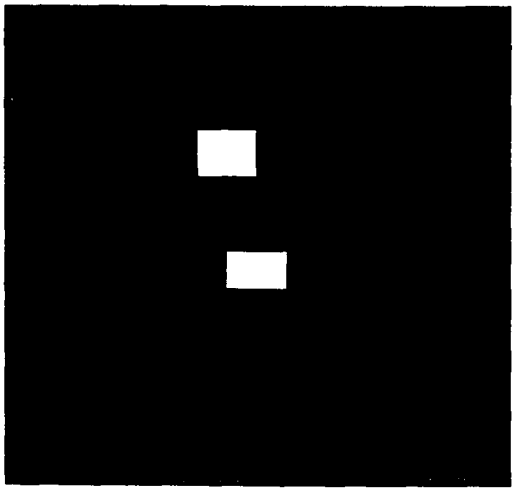


Figure 6-42 : Difference Image using images in Figures 6-40 and 6-41.

The results for the centered images are shown in Figures 6-40 and 6-41. Figure 6-42 presents the difference image obtained by differencing the images shown in Figures 6-40 and 6-41. After moving the objects to the center of the image frame, the correlation coefficient and correlation measure are computed to determine if further use of the registration algorithm is required.

Table 6-4 : Correlation coefficient and correlation measure for Figures 6-40 and 6-41.

Correlation coefficient (r)	0.919809
Correlation measure r-square	0.846048

The threshold value for any correlation depends largely on the image type. A choice of threshold values will be influenced and probably dictated by the application domain.

6.3.2.4 Case 4

In this case we address the registration of the images in Figures 6-43 and 6-44. These images are the same square, but the square in Figure 6-44 is a translation of the square in Figure 6-43. The square in Figure 6-43 is located at a corner of the image frame and this requires two iterations of the image centering program. The reason for the second centering follows.

Figure 6-43 : Image #1 Square at the corner

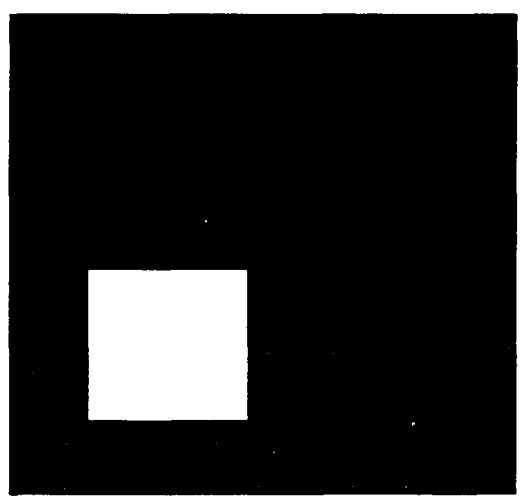


Figure 6-54 : Image # 2 Square

The wavelet analysis program is run on both images and the hull points are computed using the Wavelet Modulus Maxima Images as input and these resultant points are shown in Figures 6-45 and 6-46.

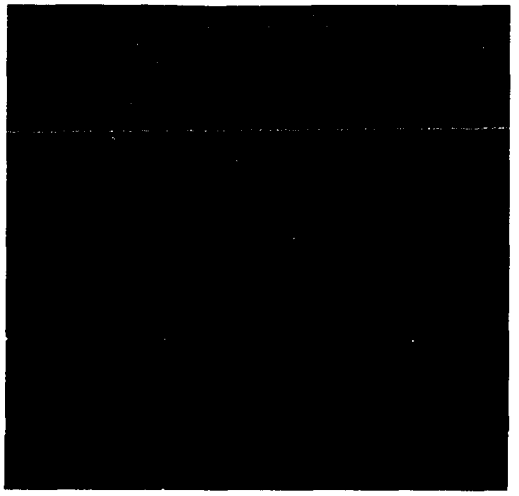


Figure 6-45 : Hull points obtained from the Modulus Maxima Image of the object in Figure 6-43.

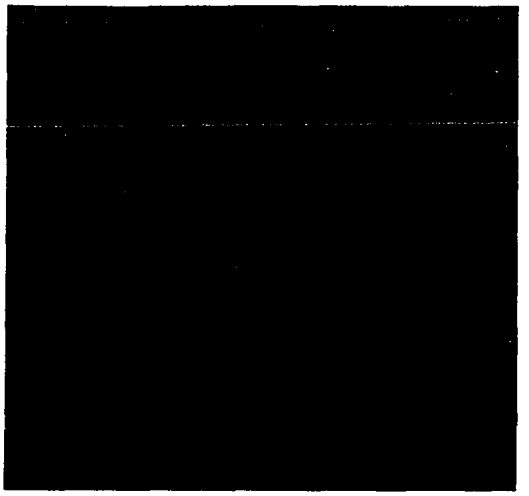


Figure 6-46 : Hull points obtained from the Modulus Maxima Image of the object in Figure 6-44.

Using these hull points, the centroid for each set of the hull points shown in Figures 6-45 and 6-46, is computed and subsequently used to compute the translation parameters. However, Figure 6-45 has only three hull points showing whereas, Figure 6-46 has five hull points showing. This occurs because one edge of the Modulus Maxima Image, for the image in Figure 6-45, was located outside the image frame.

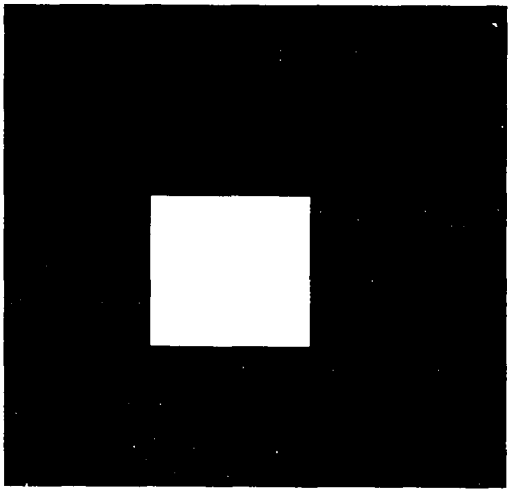


Figure 6-47 : Translated version of image in Figure 6-43 using image on Figure 6-45

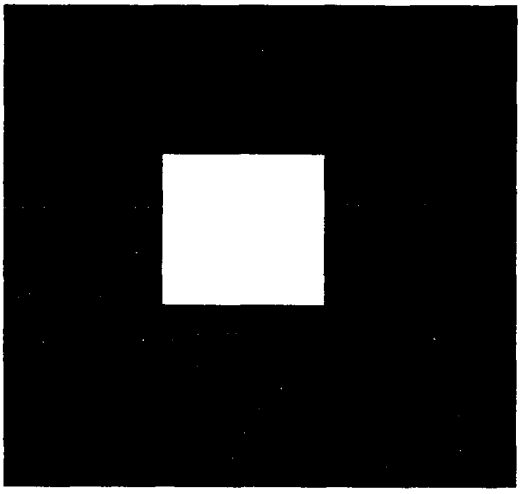


Figure 6-48 : Translated version of image in Figure 6-44 using the image in Figure 6-46.

Using the image centering algorithm the images are translated to the center of the image frame. Figure 6-47 shows the effect of this translation using the points in Figure 6-45. Figure 6-48 is the translated version of the image using the points in Figure 6-46. The difference image between the images in Figure 6-47 and 6-48 is shown in Figure 6-49.

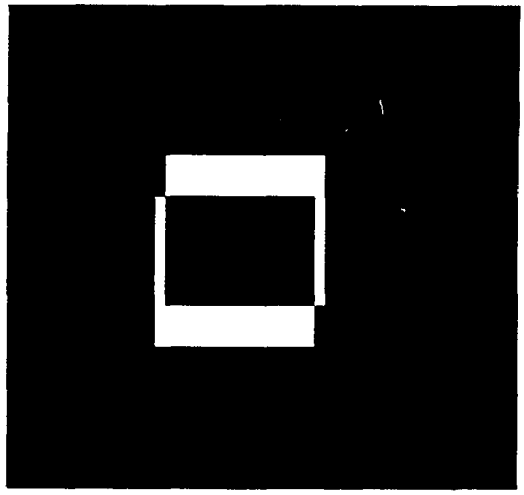


Figure 6-49 : Difference image (Figures 6-47 and 6-48 differenced).

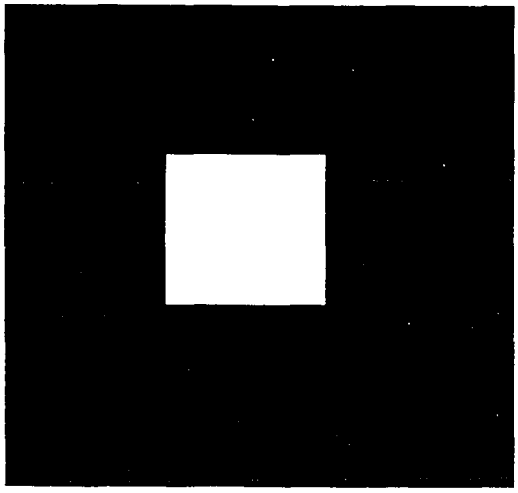


Figure 6-50 : Image in Figure 6-43 after second iteration.

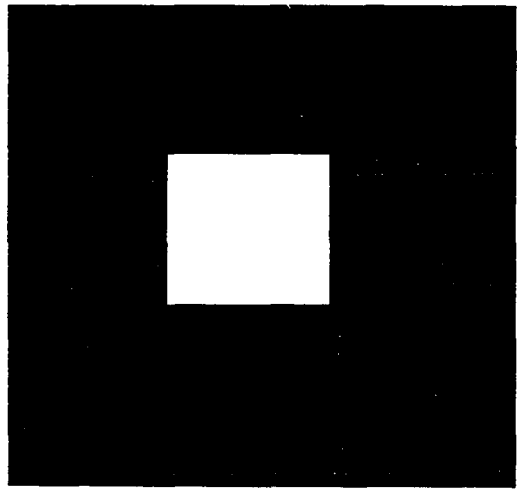


Figure 6-51 : Image in Figure 6-44 after second iteration.

Figure 6-52 : Difference Image (Image in Figures 6-50 and 6-51 differenced).

The wavelet analysis, hull computations and the image centering algorithm are rerun on the images shown in Figures 6-47 and 6-48 to produce the Figures 6-49 and 6-50. Figure 6-51 is obtained by difference the images in Figures 6-49 and 6-50.

Table 6-5 : The following table shows the correlation coefficients and measures after each iteration between the translated images.

Iteration # 1	Correlation coefficient (r) =	0.645072
	Correlation measure r-square =	0.416118
Iteration # 2	Correlation coefficient (r) =	1.000000
	Correlation measure r-square =	1.000000

6.3.2.5 Observations

1. The edges move because of convolution.
2. Our procedure appears to work for binary images in general.

6.3.3 MRI Images

We chose MRI images of the brain for additional verification and testing of our registration algorithm. Compared to the images used in section 6.3.2, MRI images are gray

scaled and have a noisy background. Again both translations and rotations of the brain MRIs are presented.

6.3.3.1 Translation

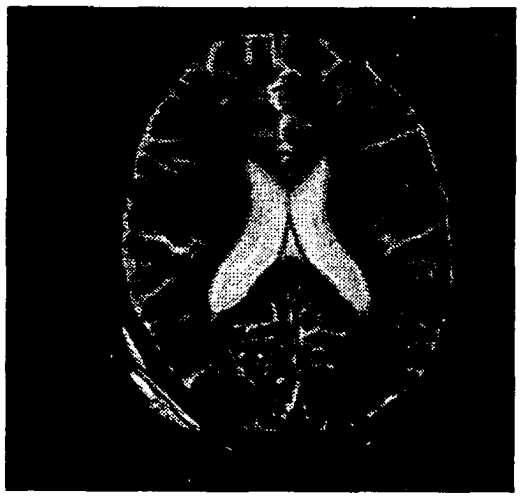


Figure 6-53 : Image #1 MRI of a brain.

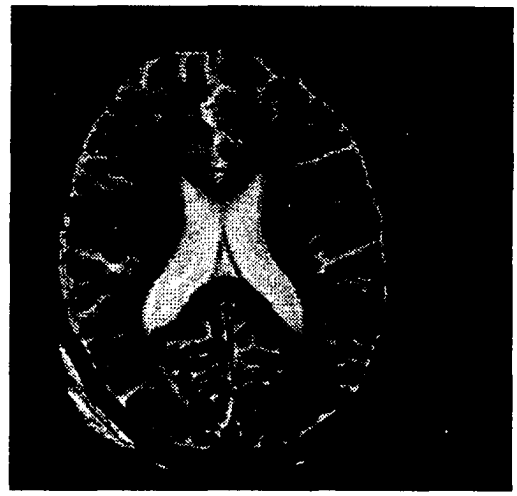


Figure 6-64 : Image # 2
Translated image of the
brain shown in Figure 6-53.

The MRI brain image in Figure 6-54 is a translation of the same MRI brain image that is shown in Figure 6-53. The images shown in Figures 6-53 and 6-54 are used to produce the difference image shown in Figure 6-55. Our registration analysis begins with the wavelet program run on both images to generate the Wavelet Modulus Maxima Image for each image shown in Figures 6-53 and 6-54. The resultant Wavelet Modulus Maxima Images for these are shown in Figures 6-56 and 6-57 respectively. The hull points computed for the Modulus Maxima Images in Figures 6-56 and 6-57 and the hull points are shown in Figures 6-58 and 6-59 respectively. The image frame centering algorithm results, to center these images, are shown in Figures 6-60 and 6-61. A comparison of the Figures 6-56 and 6-58, shows that some of the points that constitute the convex hull shown in Figure 6-58

cannot be seen in the Figure 6-56. This is an important observation because the Modulus Maxima Image serves as the input to the convex hull program.

Figure 6-55 : Difference image displayed in reverse video for clarity.

The low resolution of our viewing devices (monitor, printer paper) and the human eye's limited ability to distinguish all shades of gray, is the cause for those points to be not visible in Figure 6-56. However when the convex hull program is run, the computer uses raw gray levels and knows the existence of the points in question.

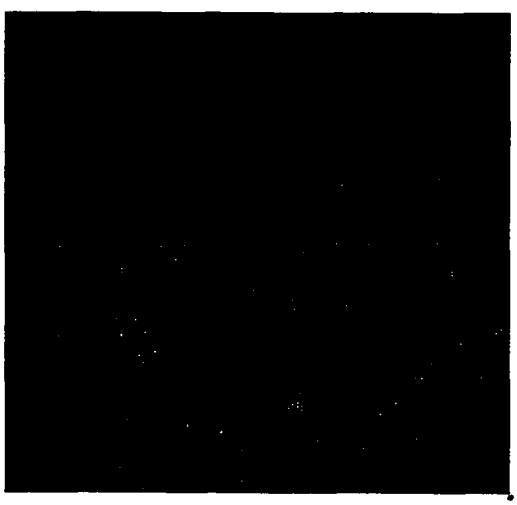


Figure 6-56 : Modulus Maxima Image for image in Figure 6-53.

Figure 6-57 : Modulus Maxima Image for image in Figure 6-54.

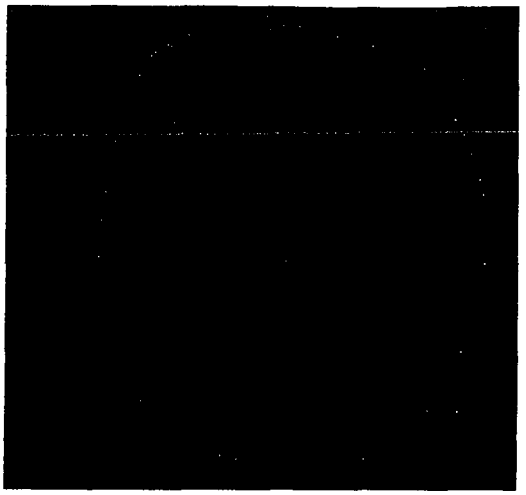


Figure 6-58: Hull points
for image in Figure 6-56.

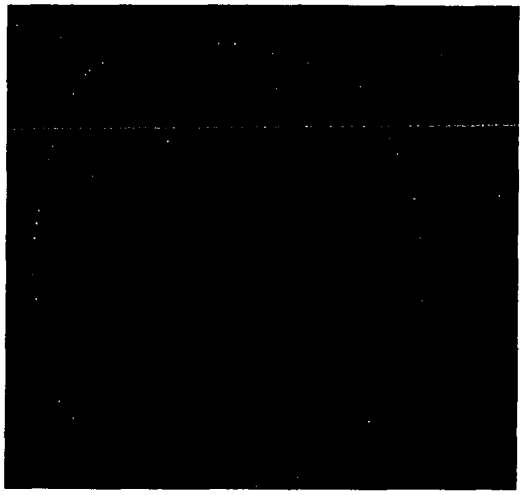


Figure 6-59 : Hull points
for image in Figure 6-57.

The convex hull program changes the gray value of the points that constitute the hull to 255, thus making them visible to the human eye. It is important to point out that our principal components analysis program, uses the location of the points on the convex hull and not the gray level.

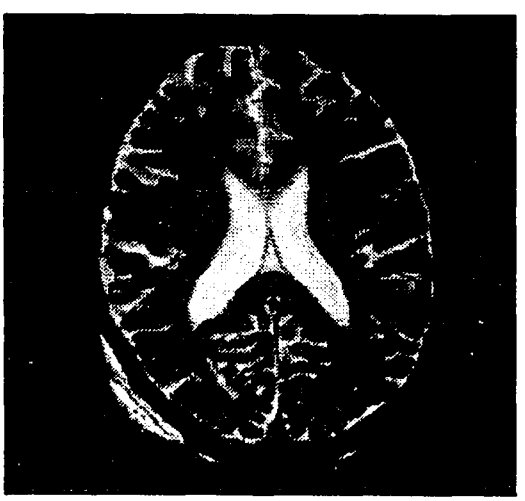


Figure 6-60 : Image in 6-53 centered
in the image frame

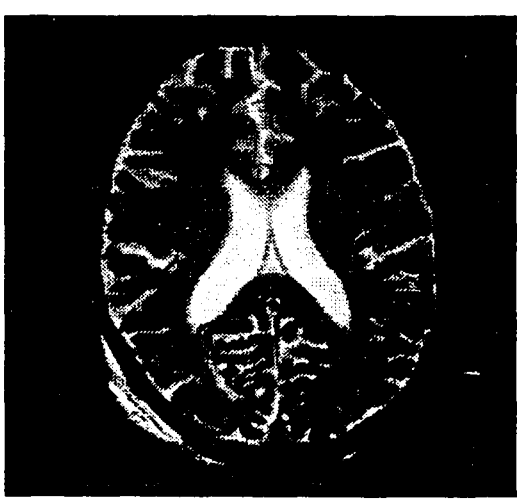


Figure 6-61 : Image in 6-54 centered
in the image frame

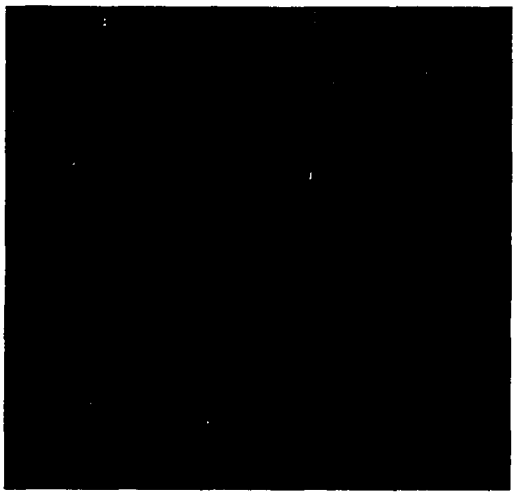


Figure 6-62 : Difference Image (Images in Figures 6-60 and 6-61 are differenced)

Table 6-6 shows the correlation coefficient and measure for the images before registration and after the centering algorithm was run. Since these correlation coefficients and measures are closer to one, the algorithm is terminated. The correlation coefficient *r* above 0.99 indicating that the accuracy of the registration is very high.

Table 6-6 : Shows the correlation coefficient and measure.

Correlation for Images in Figures 6-53 and 6-54 . (Before registration)	Correlation coefficient (r) = 0.530488 Correlation measure r-square = 0.281418
Correlation for Images in Figures 6-60 and 6-61 . (After first iteration in registration)	Correlation coefficient (r) = 0.999768 Correlation measure r-square = 0.999535

The magnitude of the translation does not affect the accuracy of the registration as long as the objects being registered remain completely within the image frame.

6.3.3.2 Rotation

This section is devoted to registration of MRI images of the brain that have been rotated. Figure 6-63 shows an MRI image of the brain and Figure 6-64 shows the same MRI rotated by 20^0. The difference image using images from Figures 6-63 and 6-64 is shown in Figure 6-65.

Figure 6-63 : Image #1 MRI of brain.

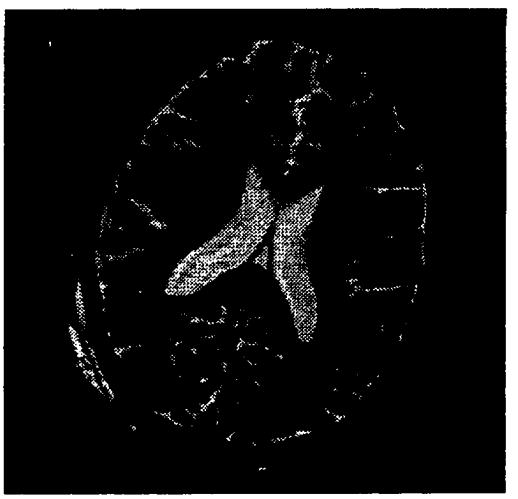


Figure 6-64 : Image # 2 (Image # 1 rotated (20^0)).

Figure 6-65 : Difference Image (Difference between Figures 6-63 and 6-64)

Figure 6-66 : Modulus Maxima Image for brain scan in Figure 6-63.

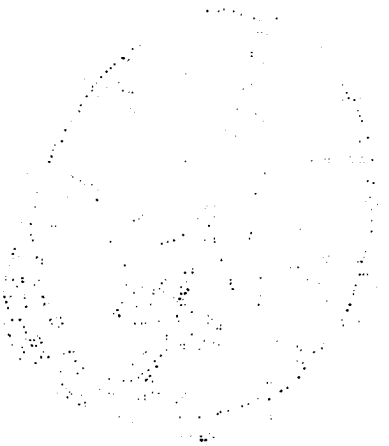


Figure 6-67 : Modulus Maxima Image for brain MRI in Figure 6-64.

The Wavelet Modulus Maxima computed for the images in Figures 6-63 and 6-64 are shown in Figures 6-66 and 6-67. The convex hull points computed from the Wavelet Modulus Maxima Images shown in Figures 6-66 and 6-67 are shown in Figures 6-68 and 6-69, respectively.

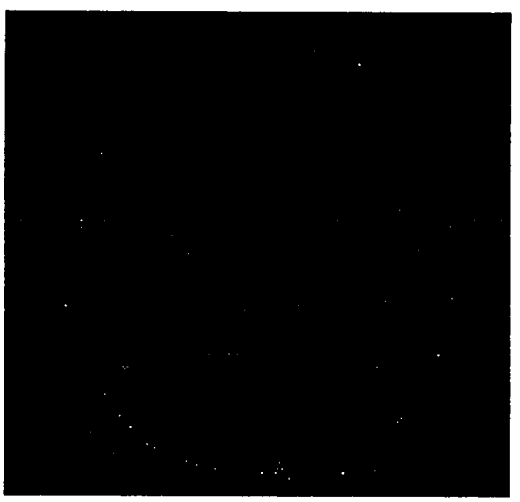


Figure 6-68 : Hull points obtained using the image in Figure 6-66 as input.

Figure 6-69 : Hull points obtained using the image in Figure 6-67 as input.

Figure 6-70 : MRI brain image shown in Figure 6-64 rotated by 24.93^0. (Counter clockwise rotation)

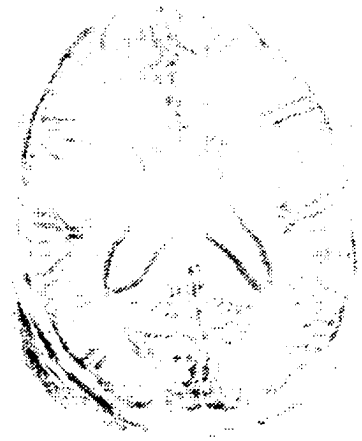


Figure 6-71 : Difference Image Images in Figures 6-63 and 6-70 differenced.

The principal component analysis uses these hull points and computes the principal axis. The principal component analysis program for the hull points shown in Figure 6-68 computes the angle the principal axis makes with the horizontal to be 80.99^0. Similarly for the hull points in Figure 6-69 the angle computes to 56.07^0. The angle of rotation for the first iteration is the angle between the principal axis from the two figures and this works out 24.93^0. The MRI brain scan shown in Figure 6-64 is rotated counter clockwise by 24.93^0 and this is shown in Figure 6-70. Figure 6-71 shows the difference between the original image (Figure 6-63) and this rotated image (Figure 6-71). The correlation coefficient between the images in Figures 6-70 and 6-63 is 0.903543. This is a significant increase from the correlation coefficient value between the unregistered images shown in Figures 6-63 and 6-64. In the next iteration the hull points are computed for the image in Figure 6-70. Using these hull points the angle the principal axis makes with the horizontal direction computes to 87.50^0. The program therefore requires that the image in Figure 6-70 be rotated by 6.51^0 in the clockwise direction. Since rotation geometrically distorts

the image with every rotation, instead of rotating the image in Figure 6-70 by 6.51^0 in the clockwise direction, we rotate the image in Figure 6-64 by $18..41^0$ in the counter clockwise direction. This new rotated image is shown in Figure 6-72. Figure 6-73 shows the difference image between the images in Figures 6-63 and 6-72. The correlation coefficient between the Figures 6-63 and 6-72 computes to 0.972718. The program then switches to the fine tuning mode because the correlation coefficient reaches above the set threshold and also because further usage of the modulus maxima technique does not lead to improvement of the correlation coefficient. The fine tuning algorithm determines that a further rotation of 1^0 in the counter clockwise direction would lead to an improvement of the correlation coefficient. Once again instead of rotating the image in Figure 6-72 by 1^0 the program achieves the same objective by rotating the image in Figure 6-64 by 19.41^0 in the counter clockwise direction. This rotated image is shown in Figure 6-74. Figure 6-75 shows the difference image between the images in Figures 6-63 and 6-74. The correlation coefficient between the Figures 6-63 and 6-74 computes to 0.991998. The algorithm terminates at this point because further movement of the image by the fine tuning routine does not lead to an increase in the correlation coefficient.

The correlation coefficients and correlation measures for this registration are shown in the table 6-7. As indicated in our algorithm (chapter 5) and as the pseudo code in section 6.3.1, the registration procedure is iterative. Iterations are continued until the correlation coefficient satisfies a threshold value, as is the case with all programs that compute solutions to problems by iteration. The threshold value used for this application domain is discussed further in section 6.4. It should be noted that only the second image is rotated, keeping the first image fixed for this and all subsequent analysis.

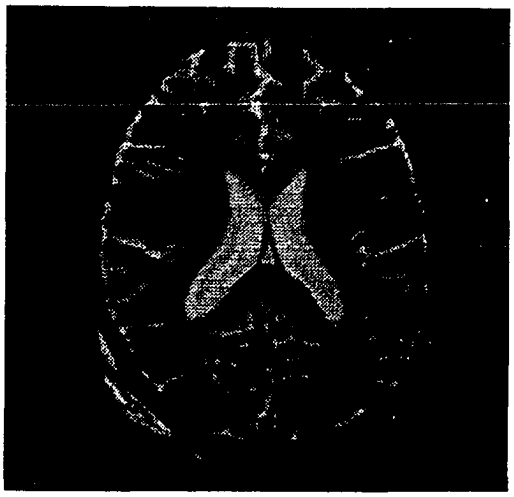


Figure 6-72 : MRI brain image shown in Figure 6-64 rotated counter clockwise by 18..41°.

Figure 6-73 : Difference Image Images in Figures 6-63 and 6-72 differenced.

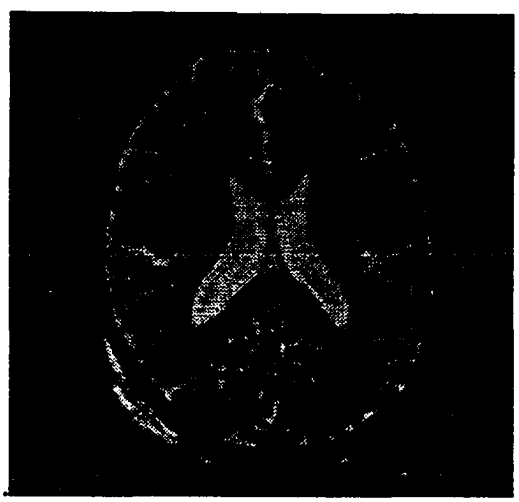


Figure 6-74 : MRI brain image shown in Figure 6-64 rotated counter clockwise by 19.41°.

Figure 6-75 : Difference Image Images in Figures 6-63 and 6-74 differenced.

Table 6-7 : This table shows the correlation coefficient and measure at the different stages of the iteration for brain MRIs.

Images being considered	**Correlation**
Images in Figure 6-63 and 6-64. (Before Registration) [Image in Figure 6-64 is rotated clockwise by 20^0]	Correlation coefficient (r) = 0.765776 Correlation measure r-square = 0.586413
Correlation between the images in Figure 6-63 and 6-70 is computed. The image in Figure 6-64 has been rotated by 24.93^0 counter clockwise to produce image in Figure 6-70.	Correlation coefficient (r) = 0.903543 Correlation measure r-square = 0.816389
Correlation between the images in Figure 6-63 and 6-72 is computed. The image in Figure 6-64 has been rotated by 18.41^0 counter clockwise to produce image in Figure 6-72.	Correlation coefficient (r) = 0.972718 Correlation measure r-square = 0.946181
Correlation between the images in Figure 6-63 and 6-74 is computed. The image in Figure 6-64 has been rotated by 19.41^0 counter clockwise to produce image in Figure 6-74.	Correlation coefficient (r) = 0.991998 Correlation measure r-square = 0.984060

Figure 6-64 is the image of the same brain shown in Figure 6-63 with a rotation of 20^0. The registration method developed in this research determines the angle to be 19.41^0.

This is very close to accurate. We once again conclude that this registration technique is fast and accurate.

Now let us consider a case where one of the MRI brain images is given a predetermined rotation of 30^0 in the clockwise direction with respect to the MRI brain image used to produce it. Our procedure converges in one iteration for this 30^0 rotation. Figures 6-76 and 6-77 are still MRI images of the brain of the same patient, but one is produced by a 30^0 rotation of the original. Figure 6-78 shows the difference image prior to registration of Figures 6-76 and 6-77. The hull points are obtained from the Wavelet Modulus Maxima Image after the wavelet analysis for Figures 6-76 and 6-77 are shown in Figures 6-79 and 6-80 respectively.

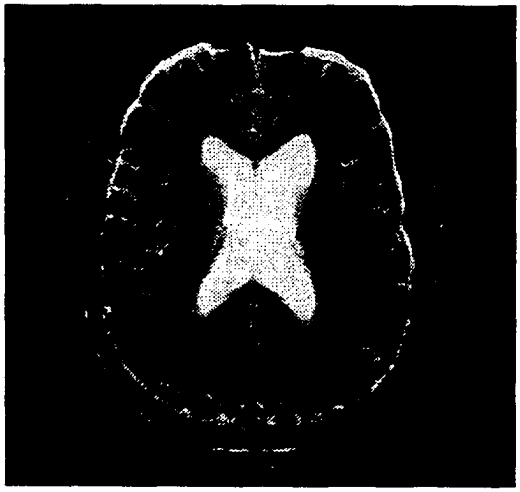


Figure 6-76 : Image #1 MRI of a brain

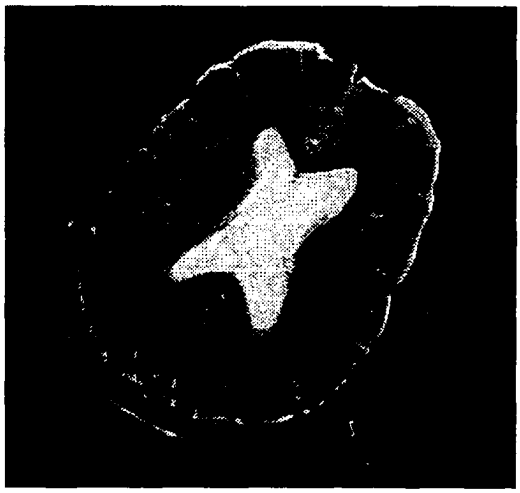


Figure 6-77 : Image # 2 Rotated (30^0)

The principal component analysis is run using the points on the convex hull of each Wavelet Modulus Maxima Image determines the angle of rotation by taking the dot product of the principal components. This angle was determined to be 29^0 in the counter clockwise direction. The resultant image after the 29^0 counter clockwise rotation is shown in Figure 6-81. The image in Figure 6-81 is subtracted from the original image in Figure

6-76 and this is shown in Figure 6-82. Table 6-8 shows the relationship between the images and the correlation coefficient and measure for this case. The correlation coefficient is above 0.98 indicating that the registration procedure does work accurately.

Figure 6-78 : Difference Image (Images in Figures 6-76 and 6-77 are differenced)

Figure 6-79 : Hull points obtained for the image in Figure 6-76.

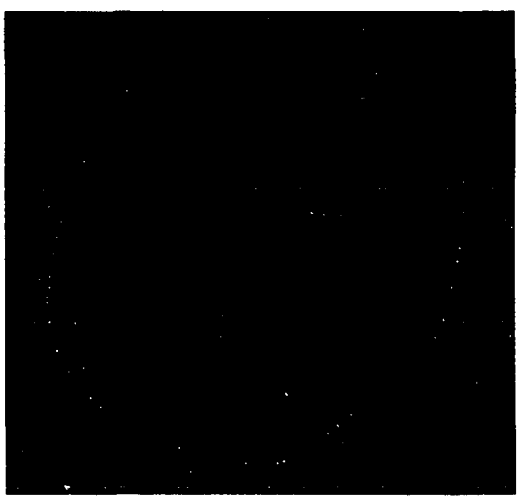


Figure 6-86 : Hull points obtained for the image in Figure 6-77.

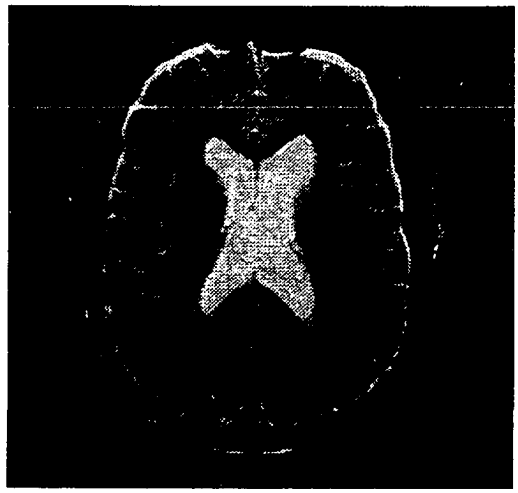


Figure 6-81 : MRI brain image shown in figure 6-77 rotated by 29^0 (counter clockwise rotation).

Figure 6-82 : Difference Image
Images in figures 6-82 and 6-81 differenced.

Table 6-8 : The behavior of correlation coefficient and measure as the images are rotated.

Images being considered	**Correlation**
Images in Figures 6-76 and 6-77. Image in Figure 6-76 is rotated clockwise by 30^0 to produce image in Figure 6-77.	Correlation coefficient (r) = 0.756488 Correlation measure r-square = 0.572273
Images in Figures 6-76 and 6-81. The image in Figure 6-77 has been rotated by 29^0 counter clockwise to produce image in Figure 6-81.	Correlation coefficient (r) = 0.984874 Correlation measure r-square = 0.969976

6.3.3.3 Rotation, Translation and Artifacts

In this section the case of an MRI image that has undergone translation and rotation is presented. In addition in the rotated MRI image of the brain artifacts have been introduced. This has been done to simulate internal changes in the MRI of the brain.

Regardless of the extent and shape of the internal changes the procedure developed in this research is able to accurately register the images. Figures 6-83 and 6-84 represent the initial two images to be registered. Figure 6-85 is obtained by taking the difference between the images in Figures 6-83 and 6-84.

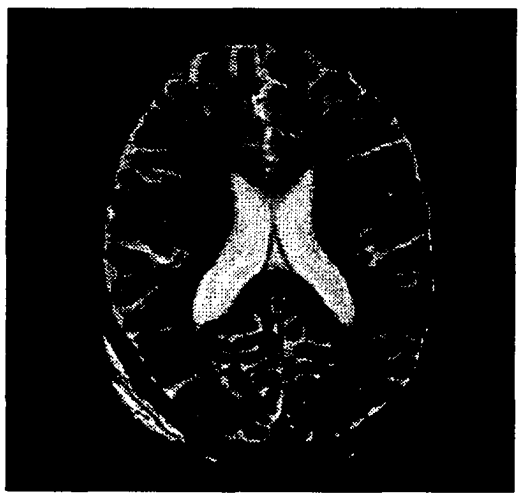


Figure 6-83 : MRI of the brain (IMAGE# 1)

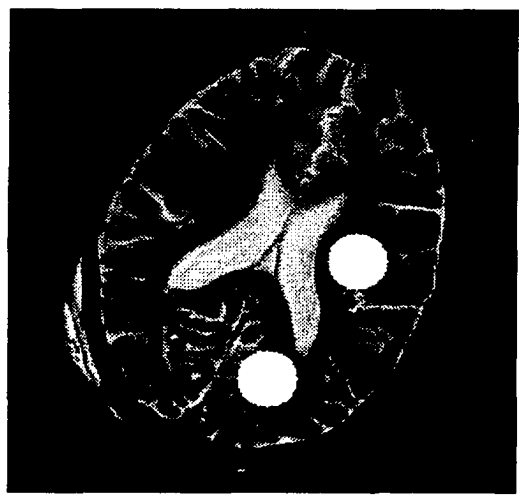


Figure 6-84 : The same MRI as in Figure 6-83 rotated by 30^{0} artifacts added.

The registration procedure is applied to the initial images and after a few iterations we get the registered images shown in Figure 6-86. The image in Figure 6-86 is obtained after applying the transformations to the Figure in image 6-84. Figure 6-87 shows the difference between the image in Figures 6-83 and 6-86. The initial correlation between the images to be registered (Figures 6-83 and 6-84) is 0.688755. The correlation coefficient between the registered images shown in Figures 6-83 and 6-86 is 0.900972. The correlation coefficient is not higher in this case partially due to misalignment and partially due to the images being registered having dissimilarities. Here in lies one of the deficiencies in using the correlation coefficient to access how well two images registered. It is likely that two brain images with a great amount of internal differences but with perfect spatial alignment could have a

lower correlation coefficient. Table 6-9 shows the correlation coefficient between the images before and after registration.

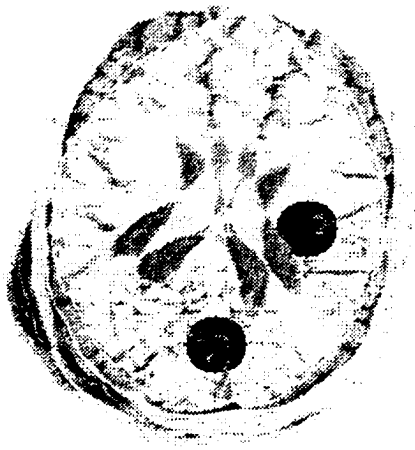


Figure 6-85 : This figure shows the difference between the images in Figures 6-83 and 6-84.

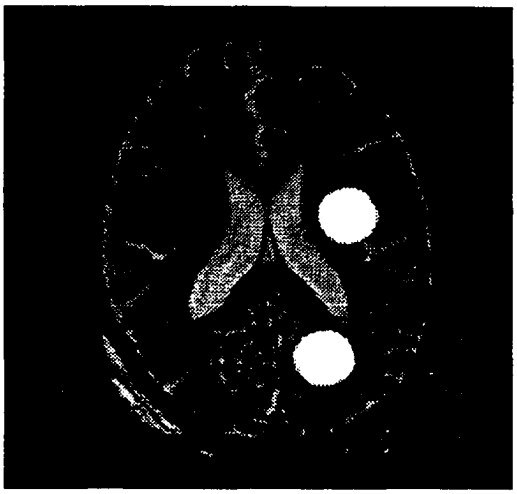


Figure 6-86 : Image 6-84 after registration.

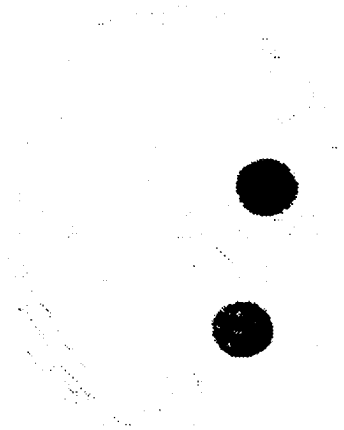


Figure 6-87 : Difference image obtained by taking the difference between the images in Figures 6-83 and 6-86.

Table 6-9 : The behavior of correlation coefficient for the case when artifacts are added to a rotated MRI of the brain.

Images being considered	**Correlation**
Images in Figures 6-83 and 6-84. (Unregistered Images)	Correlation coefficient (r) = 0.688755 Correlation measure r-square = 0.474384
Images in Figures 6-83 and 6-86. (Registered Images)	Correlation coefficient (r) = 0.900972 Correlation measure r-square = 0.811751

The registration procedure computes the angle of rotation to be 29.6^0. This is very close to 30^0 the actual angle of rotation indicating that the procedure is very accurate.

6.4 Observations and conclusions

Several observations about this registration algorithm and procedure follow:

1. Appendix B offers an explanation for changes in the gray scale distribution once the images are rotated. The experimental results presented in part of Appendix B show a sample distribution of gray levels for a square that had all its pixels with gray value equal to a 100 and was then subjected to various rotations. Because of the changes in the gray scale distribution even after registration, the difference image will have some pixels at different gray levels, from the rotation.
2. The number of iterations shows no apparent correlation with the degree of rotation. A number of angles of rotation were chosen ranging from 2^0 to 90^0 for testing. In all the cases tested, the algorithm converged. Additionally using different images with the same angle of rotation could even converge with a different number of iterations. So it

is not reasonable to relate the angle of rotation with the number of iterations required for convergence because this also depends on the images.

3. For gray scale images, like the MRI, which have noise as part of the image frame, some denoising is required. In this algorithm, denoising is done by ignoring all wavelet coefficients in the Wavelet Modulus Maxima Image that are less than 40. The maximum value for these coefficients is 255. The threshold value may need to be set at different values for different applications.
4. When is an image is registered, the correlation coefficient and measure is examined and used to terminate the registration algorithm. In all cases tested, this has worked successfully. However, one must be cautious with the values used for the threshold because of the possibility that brain images, with significant internal changes, can yield a lower correlation coefficient even if registered with a previous scan of the same brain which did not have these internal changes.
5. We found that the edges located by the wavelet program moved. It was the convolution that caused movement of the edges. Convolution is part of the normal wavelet analysis and synthesis. This chapter provides equations for bounds on the movement of the edges for each step in the wavelet iteration program.
6. The algorithm affords a choice of the size of the Wavelet Modulus Maxima Image to use for computing the hull and on the translation and rotation parameters. The trade-off is the smaller the image size, the smaller the number of hull points produced and this affects the accuracy. At most, the wavelet program should be limited to three levels of iteration.
7. The time complexity of the Wavelet Modulus Maxima algorithm is $O(n^2)$. The principal component analysis algorithm also has a time complexity of the order of

$O(n^2)$. However, since in each stage of the wavelet analysis the size of the image is reduced by a factor of 2^{stage}, where stage denotes the wavelet iteration routine, the time complexity for the principal component analysis is actually $O((2^{stage}n)^2)$. The hull algorithm time complexity is of $O(n_1 \log n_1)$ where n_1 is the size (the input to the hull program) of the Wavelet Modulus Maxima Image. This Wavelet Modulus Maxima Image matrix is a very sparse matrix and of size $2^{stage}n$, so the time complexity of the hull program can be more accurately written to be of the order $O(2^{stage}n \log(2^{stage}n))$. This is true when downsampling is used as part of the Wavelet Analysis program - we get a smaller downsampled image with each iteration. However, in the version of the algorithm the image frame size stays constant, but the filter size is incremented in each iteration by upsampling it with zero. The Modulus Maxima Image contains far fewer points regardless of the algorithm used. In the actual coding of the program we use a doubly linked list to store the points on the Modulus Maxima Image and this list serves as input to the program that computes the convex hull.

6.4.1 Limitations and Failures

The technique does not work for the following situations:

1. Busy images which contain many objects probably will not be correctly registered by this technique. This is because the image centering algorithm is designed to center one object in the image; typically true in medical images.
2. The images must be completely contained within the image frame. Rotation should not yield parts of the object falling outside the image frame. The algorithm does not work for aerial surveillance images but with some modification can be adapted to work with these images. In general, the techniques can be used in this application domain of

Medical Imaging but the algorithms must be modified to work with aerial surveillance images.

3. This algorithm does not work when dissimilar objects are registered and the results are unpredictable. The registration technique does not predict the dissimilarity of objects. But, it can serve to register similar objects that lie completely within the image frame. The procedure has been tested for only one object in the image frame; typical for most medical images.
4. Noisy images tend to produce inaccurate results. However, the noise level in most MRI images is at a level where a cutoff value for the wavelet coefficients, denoises the image to a level where this registration performs well. For MRI images a cutoff of 60 denoises the images well.
5. Gray level redistribution is caused by the rotation and how we interpolate pixel intensities at integral pixel locations after rotation. The gray level in the rotated image at any point, gets its value from one or more adjacent pixels in the unrotated image (see Appendix B). Appendix B also presents the redistribution of gray values when a simple rectangle is rotated.
6. Though the techniques are more general, this algorithm was specifically designed to automatically register medical images which are MRI scans of the brain. To make the technique applicable to other areas, the algorithm may have to be modified. Threshold cutoff values on wavelet coefficients must be adjusted to the image types being registered. This is a limitation due to using correlation as a measure to estimate the accuracy of registration.
7. The number of iterations is unpredictable and cannot be correlated to the angle of rotation. The number of iterations depends upon the MRI scans being registered. The

procedure is iterative because the gray level redistribution caused by rotation leads to the object boundaries being detected with a degree of error. Additionally the detection of edges using wavelets adds to this error. Therefore the principal component analysis yields the principal axis for the images being registered, with a margin of error that requires the procedure to be iterative.

CHAPTER 7

SUMMARY AND FUTURE DIRECTIONS

Registration of images is an important area of image processing. Any time a comparison is made between images, as in the case of medical images, aerial surveillance photos to predict changes in the terrain, or automatic target recognition; these images have to be registered. Registration is an important preprocessing step in object recognition and object classification. As the world gets more automated, the need to register images automatically will be more necessary. In medical imaging, medical images have to be registered for doing atlas studies, for obtaining diagnostic information, and for other clinical purposes. Registration of images has been taken for granted in many situations. It is for this reason that the state-of-the-art in this area is mostly art. There is no standardization or even agreement on validation techniques at this time. The age old technique of "eye-balling" has been used as a technique to register images. Because some gray level changes cannot be discerned with the human eye, what constitutes registered images, has been variable and operator dependent. Depending on the resolution of the image, it is almost impossible to point out neighboring pixels with a mouse pointer. The methodology developed in this dissertation addresses this problem. A quantitative measure of knowing how well images are registered by using the correlation coefficient is presented. The correlation coefficient has deficiencies and affects all registration techniques developed. Rotation affects all registration techniques since it distorts the object. Because of the gray level redistribution, the detection of edge boundaries is affected. The number of iterations to register images is unpredictable, but after the iterations it does register the objects.

7.1 Contributions

The contributions of this dissertation are:

1. A new method to register images is proposed. This method is very fast and accurate. Our method uses a smaller and more sparse matrix to compute the translation and rotation parameters and thus contributes an algorithm with improved time complexity while keeping the accuracy to quantified tolerable levels.
2. It automates the process of registration, which is mostly a manual procedure in medical imaging at this time.
3. The procedure uses wavelets to register images and thus may be more acceptable as a preprocessing step to wavelet based compression techniques; assuming wavelet compression techniques become more acceptable in medical imaging.
4. This dissertation shows the relationship of the scatter plot to image registration. The suggestion is that it may be a better technique for "eye-balling" purposes than looking at raw images. As similar images are better registered the scatter plot clusters about the 45^0 line passing through the origin.
5. This dissertation strongly suggests using the correlation coefficient to measure the extent to which images are registered. Since a correlation coefficient may fail at times to obtain the threshold, "eye-balling" of the scatter plot could be added. If the method is made semi-automatic by using a trained operator to assist in the validation of the convergence, this method becomes more accurate.
6. This method makes a contribution in detecting and documenting the location and the movement of edges when wavelet techniques are used. The dissertation provides a formula that predicts the location of edges after each iteration.

7. Rotation of images introduces the need for interpolation to get smoother edges. Pixels are located at integral positions; however, during rotation some pixels in the rotated image do not fall at integral coordinate positions. Appendix B documents the effects of truncating floating point values to integral values. Appendix B also documents the impact of using bilinear extrapolation. The discovery of this redistribution of gray values is another contribution. The effect of a redistribution of gray values is significant when rotation is part of any technique that does lossless compression. This redistribution is not noticeable when rotated images are visually inspected. Computer based techniques are affected by this redistribution. In this research we have used bilinear interpolation with backward mapping. Bilinear interpolation with forward mapping as well as other interpolation techniques used during rotation will result in a different redistribution of gray levels thus affecting any further analysis that is made using the result; for example the computation of principal components, correlation coefficient, the location of edges, etc. However even for a different choice of the interpolation technique, this registration method will work.
8. This research shows that no registration procedure can accurately register images by detecting edges if rotation is involved. Rotation essentially redistributes gray values causing edges to be detected with a margin of error. So even the age old technique of "eye-balling" will have errors because a picture with redistributed gray levels will lead even trained operators to detect edges with some error. Any validation technique that uses gray levels will also have some built in inaccuracy.

7.2 Future Directions

This wavelet based registration technique is desensitized to the internal changes of the object being registered. However, the technique could be more application domain

sensitive if it is made only partially sensitive to internal features of the objects being registered. The internal features that the procedure is made sensitive to should be dictated by the objects being registered. This technique could be improved to register busy images and images with multiple objects. A better denoising algorithm would help in detecting true edges as compared to detecting some edges that are due to noise (this is some times referred to as a false positive). The cutoff threshold must be determined by using training images from the application domain. This technique would be improved if the number of iterations were more predictable.

Registration can be extended to automatic target recognition and data mining. In this dissertation research, we used two dimensional images but the technique can be extended to three dimensional volume data sets. Future research may also consider adjusting for edge movements and better edge detection. An extension that is being pursued is to execute this code in a parallel computing environment.

BIBLIOGRAPHY

[Akl78] Akl S. G. and Toussaint G. T., "Efficient convex hull algorithm for pattern recognition applications", Proceedings of the 4th International Joint Conference on Pattern Recognition, Kyoto, Japan, 1978.

[Akl79] Akl S. G., "Two remarks on a convex hull algorithm", Information Processing Letters, 8, 1979.

[Alpert90] Alpert N. M., Bradshaw J. F., Kennedy D. , Correia J. A., "The Principal Axes Transformation - A Method of Image Registration", The Journal of Nuclear Medicine, Vol. 31, No 10, October 1990.

[Anton95] Anton H., and Rorres C., "Elementary Linear Algebra", John Wiley & Sons, New York, 1994.

[Aravena95] Aravena G., "Lecture Notes - Course on Wavelets", Department of Electrical Engineering, Louisiana State University, July 1995.

[Barnea72] Barnea D. I., Silverman, H. .,"A Class of Algorithms for Fast Digital Image Registration", IEEE Transaction on Computers, Vol. c-21, No 2, February 1972.

[Bracewell86] Bracewell R., "The Fourier Transform and its Applications", McGraw-Hill, New York, 1986.

[Brooks89] Brooks, R. A., "Symbolic reasoning among 3-D models and 2-D images", Technical Report AIM.-345, Standford Artificial Intelligence Lab, June 1981, cross-reference from "Matching Images Using Linear Features", Gerard Medioni and Ramakant Nevatia, IEEE, PAMI, Nov 89.

[Brown92] Brown G. L., "A Survey of Image Registration Techniques", ACM Computing Surveys, Vol. 24, No 4, 1992.

[Bruce96] Bruce, A., Donoho, D. and Gao H., "Wavelet Analysis", IEEE Spectrum, Oct. 1996.

[Burden93] Burden, L. R. and Faires, J. D., "Numerical Analysis", PWS Publishing Company, Boston, 1993.

[Cappelletti89] Cappelletti, J. D. and Rosenfeld, A., "Three-Dimensional Boundary Following", Computer Vision, Graphics, and Image Processing, 48, 1989.

[Castleman96] Castleman, K., "Digital Image Processing", Prentice Hall, Englewood Cliffs, New Jersey, 1996.

[Chatfield80] Chatfield C. and Collins A., "Introduction to Multivariate Analysis" Chapman and Hall Press, London, 1980.

[Chan93] Chan, Y. T., "Wavelet Basics", Kluwer Academic Publishing, Boston, 1993.

[Cideciyan95] Cideciyan, A. V., "Registration of Ocular Fundus Images", IEEE Engineering in Medicine and Biology, January / February 1995.

[Clarysse91] Clarysse P., Gibon D., Rousseau J., Blond S., Vasseur C., "A computer assisted system for 3-D frameless localization of stereotaxic MRI", IEEE Transaction, Medical Imaging, 1991.

[Clasen80] Clasen, T. A. C. M. and Mecklenbrauker W. F. G. , "The Wigner distribution - a tool for time-frequency signal analysis. Part 1: Continuous-time signals", Philips Journal of research., vol. 35, # 3, 1980

[Cohen89] Cohen A. and Kovacevic, J., "Wavelets : The Mathematical Background", Proceedings of the IEEE, Vol. 84, No 4, April 1996.

[Cooley71] Cooley W. W., "Multivaraiate Data Analysis", John Wiley and Sons, Inc. New Yrk, 1971.

[Daubechies93] Daubechies I., "Different Perspectives on Wavelets", Proceedings of Symposia in Applied Mathematics, Volume 47, American Mathematical Society, January 1993.

[Daubechies88] Daubechies, I., "Orthonormal bases of Compactly supported Wavelets", Communication of Pure and Applied Math., vol. 41, Nov 1988.

[Dougherty87] Dougherty E. R. and Giardina C. R., "Matrix Structured Image Processing", Prentice-Hall, Inc. New Jersey, 1987.

[Duchowski98] Duchowski, A. T. , "3D Wavelet Analysis of eye movements", Proceedings of the SPIE, Conference on AeroSense, Aerospace / Defense Sensing and Controls, Orlando, Florida, April 1998.

[Edwards94] Edwards C. H. and Penney D. E., "Calculus with Analytic Geometry", Prentice Hall, Englewood Cliffs, New jersey, 1994.

[Elsen91] Van Den Elsen P.A., Viergever M.A., "Marker guided registration of Electromagnetic dipole data ith tomographic images", cross-referenced in "Information Processing in Medical Imaging", editors: Colchester A.C. F., and Hawkes D. J., Springer-Verlag, 1991.

[Elsen93] Elsen, Petra A. van den, Pol, Evert-Jan D., and Viergever, Max A., "Medical Image Matching - A Review with Classification", IEEE Engineering in Medicine and Biology, March 1993.

[Evans89] Evans A. C., Marrett S., Collins L., Peters T. M., "Anatomical functional correlative analyis of the human brain using three dimensional imaging system", Proceedings SPIE, vol 1092, Medical Imaging III: Image Processing, SPIE Press, Bellingham, WA, 1988.

[Faber91] Faber T. L., McColl R. W., Opperman R. M., Corbett J. R., Peshoch R. M., "Spatial and temporal registration of Cardiac SPECT and MR Images: Methods and evaluation, Radiology, 179(3), 1991.

[Faires93] Faires D. J., and Burden, Richard L., "Numerical Analysis", PWS Publishing Company Boston, 1993.

[Fournier94] Fournier, A., "Siggraph 94 Course Notes on Wavelets", Siggraph 1994.

[Gabor46] Gabor D., "Theory of Communications", Journal Intitute of Electrical Engineering, vol. 93, 1946.

[Greitz80] Greitz T., Bergstrom M., Boethius J., Kingsley D., Ribbe T., "Head fixation system for integration of radiodiagnostic and therapeutic procedures", Neuroradiology, # 19, 1980.

[Herbin89] Herbin, M., Venot, A., Devaux, J.Y., Walter, E., Lebruchec, J.F., Dubertret, L., and Roucayrol, J.C., "Automated Registration of Dissimilar Images: Application of Medical Imagery", Computer Vision, Graphics, and Image Processing, 48, 1989.

[Hill91] Hill D. L. G., Hawkes D. J., Crossman J. E., Gleeson M. J., Cox T. C. S., "Registration of MRI and CT images for skull base surgery using point like anatomical features", British Journal of Radiology, # 69, 1991.

[Holshneider93] Holshneider M., Kronland-Martinet R., Morlet J. and Tchamitchian P., "A real-time algorithm for signal analysis with the help of the wavelet transform", cross reference from Stephane Mallat, et. al "Singularity Detection and Processing with Wavelets", 1993.

[Hotelling33] Hotelling H., "Analysis of a complex of statistical varaibles into principal components", The Journal of Educational Psychology, Vol 24, 1933.

[Kadambe92] Kadambe, S., "A Comparision of the Existence of 'Cross Terms' in the Wigner Distribution and the Squared Magnitude of the Wavelet Transform and the Short Time Fourier Transform", IEEE Transactions on Signal Processing, Vol 40, #10, October 1992.

[Kaiser94] Kaiser G., "A Friendly Guide to Wavelets", Birkhauser, Boston, 1994.

[Karadimitriou96] Karadimitriou K., "Set Redundancy, the enhanced compression model, and methods for compressing sets of similar images, Ph.D. Thesis, Louisiana State University, Baton Rouge, Louisiana, 1996

[Karadimitriou97] Karadimitriu K. and Tyler, J., "Set Redundancy and image compression in Medical Image databases, IEEE Transactions in Medical Imaging, - under review, 1997.

[Kapouleas91] Kapouleas I., Alavi, A. A., Wayne M., Gur, R. E., and Weiss, D. W., "Registration of Three-dimensional MR and PET Images of the Human Brain without Markers", Radiology, December 1991.

[Leithhold92] Leithhold L., "The Calculus with Analytic geometry", Harper and Row Publishers, New York, 1992.

[Lemoine91] Lemoine D., Barillot C., Giband B., Pasqualini E., "An anatomical based 3-D registration of multi-modality and atlas data in neurosurgery" cross-referenced in "Information Processing in Medical Imaging", editors: Colchester A.C. F., and Hawkes D. J., Springer-Verlag, 1991.

[Lu94] Lu J., Healy Jr., D. M., and Weaver, J. B., "Contrast enhancement of medical images using multiscale edge representation", Optical Engineering, Vol 33 No 7, July 1994.

[Lu96] Lu J., "Image Deblocking via Multiscale Edge Processing", Proceedings of the SPIE International Conference on Wavelets and Mathematical Imaging, Denvar, August 1996.

[Mallat89] Mallat G. S., "Multifrequency Channel Decompositions of Images and Wavelet Models", IEEE Transactions on Acoustics Speech and Signal Processing, Vol. 37, # 12, December 1989.

[Mallat92] Mallat, S., and Hwang, W.L., "Singularity Detection and Processing with Wavelets", IEEE Transaction on Information Theory, Vol. 38, # 2, March 1992.

[Mallat92] Mallat S. G. and Zhong, S., "Characterization of Signals from Multiscale Edges", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No 7, July 1992.

[Mallat97] Mallat S. G., "Wavelets", Proceedings of the SPIE, Conference on AeroSense, Aerospace / Defense Sensing and Controls, Orlando, Florida, April 1997.

[Medioni84] Medioni, G. and Nevatia, R., "Matching Images Using Linear Features", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. Pami-6, No 6, November 1984.

[Meyer95] Meyer, C. R., Leichtman, G. S, Brunberg, J. A., Wahl, R. L., Quint, L. E., "Simultaneous Usage of Homologous Points, Lines and Planes for Optimal 3-D Linear Registration of Multimodality Imaging Data", IEEE Transactions on Medical Imaging, Vol. 14, No 1, March 1995.

[Meyer92] Meyer, Y., "Wavelet and Operators", Cambridge University Press, 1992.

[Orr96] Orr, R. S., and Gandhi, P. P., "Wavelets in Communication Systems", Workshop, IEEE International Communications Conference, Dallas, June 1996.

[Oppenheim89] Oppenheim, A. V., and Schafer, R. W., "Discrete-Time Signal Processing", Prentice Hall Englewood Cliffs, New Jersey, 1989.

[Pohl84] Pohl I., "A Book on C", Benjamin/Cummings Publishing Company, Inc. California, 1984.

[Prasad97] Prasad L. and Iyengar S. S., ""Wavelet Analysis with Applications to Image Processing", Prentice Hall Englewood Cliffs, New Jersey, 1997.

[Proakis95] Proakis, J. G. and Manolakis, D. G., "Digital Signal Processing", Prentice_hall of India, New Delhi, 1995.

[Rabbani91] Rabbani M., and Jones P. W., "Digital Image Compression Techniques", SPIE Press, Bellingham, Washington, USA, 1991.

[Rasmussen93] Rasmussen, H. O., " The Wavelet Gibbs Phenomenon", in "Wavelets, Fractals and Fourier Transforms" edited by Farge, M., Hunt J. C. R. and Vassilicos J. C., Clarendon Press, 1993.

[Rioul91] Rioul, O. and Vitterli, M., "Wavelets and Signal Processing", IEEE Signal Processing Magazine, October 1991

[Rioul92] Rioul O. and Flandrin P., "Time-Scale Energy Distributions: A General Class Extending Wavelet Transforms", IEEE Transactions on Signal Processing, Vol. 40, #2, July 1992.

[Rousseau92] Rousseau J., Clarysse P., Gibon D., Blond S., Bradai N., "A frameless method for 3-D MRI and CT-guided stereotaxic localisation", European Radiology, No 2(1), 1992.

[Schad87] Schad L. R., Boeseche R., Schlegel, W., Hartmann G. H., Strum V., "Three dimensional image correlation of CT, MR and PET studies in radiotherapy treatment and planning of brain tumors", Journal of Computer Assisted Tomography, 11(6), 1987.

[Sharman97.1] Sharman R., Tyler J. M. and Oleg P., "Registration of Medical Images on Massively Parallel machines", ", Proceedings of the SPIE, Annual Conference, Application of Digital Image Processing XX, Optical Science, Engineering and Instrumentation, July 1997. .

[Sharman97.2] Sharman R., and Tyler, John M., "Wavelet based registration and compression of Sets of Images", Proceedings of the SPIE, Conference on AeroSense, Aerospace / Defense sensing and Controls, Orlando, Florida, April 1997.

[Sharman96] Sharman R., and Tyler, J., "Wavelet Based Registration of Medical Images", Conference on Wavelets and Operators, sponsored by UNCC, AFSOR and NSF, UNCC, Charlotte, July 1996.

[Strang96] Strang, G. and Nguyen T., "Wavelets and Filter Banks", Wellesley Cambridge Press, Wellesley, MA 02181, 1996.

[Thevenaz97] Thevenaz P. and Unser M., " Spline Pyramids for Inter-Modal Image Registration Using Mutual Information", Proceeding of SPIE Conference on Wavelet Applications in Image and Signal Processing V, San Diego California, 1997.

[Toennies88] Toennies, K. D., Udupa, J. K., Wornom III, I. L., and Buchman, S. R., "Registration of 3D Objects and Surfaces", Medical Imaging, IEEE, EMBS, 1988.

[Unser97] Unser M. A., "Ten good reasons for using Spline Wavelets", Proceeding of SPIE Conference on Wavelet Applications in Image and Signal Processing V, San Diego California, 1997.

[Vitterli92] Vitterli, M. and Herley, C., "Wavelets and Filter Banks : Theory and Design", IEEE Transactions on Signal Processing, Vol. 40, # 9, September 1992.

[Weaver83] Weaver, J. H., "Application of Discrete and Continuous Fourier Analysis", John Wiley & Sons, New York, 1983.

[Weeks97] Weeks, A. R., "Fundamentals of Electronic Image Processing", Tutorial, SPIE 11th Annual International Symposium on Aerosense, Orlando, Florida, 1997.

[Wigner32] Wigner, E, "On the quantum correction for thermodynamic equilibrium", Physics Review., vol. 40, 1932.

[Wolberg90] Wolberg, G., "Digital Image Warping", IEEE Computer Science Press, Los Alamitos, CA, 1990.

[Young93] Young, R., "Wavelet Theory and its Application" Kluwer Academic Publishing, Boston, 1993.

[Zhong90] Zhong S., "Edges representation from Wavelet Transform maxima", Ph.D. Thesis, New York University, September 1990.

APPENDIX A

CONVEX HULL

A convex hull was used in determining the principal axis and the required rotational parameters. In this Appendix, the convex hull algorithm that was modified, for this dissertation is detailed.

The convex hull $CH(S)$ of a set of points S is by definition the smallest convex set[1] containing S (i.e., the convex polygon with the smallest area that includes all the points of S, either as its vertices or as interior points [Akl93]). We specify $CH(S)$) by listing the points of S which are on the boundary of $CH(S)$, in clockwise order.

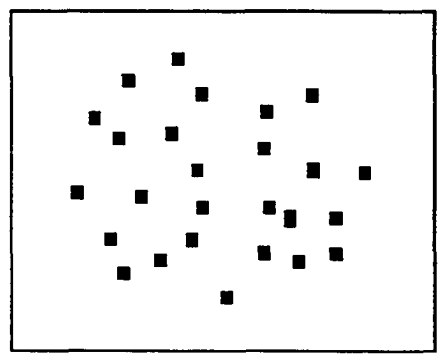


Figure A-1 : A set S of points, indicated by the small squares in the figure

[1] A set is convex if and only if the line segment determined by every pair of its points lies entirely in the set.

The convex hull for the set of points in Figure A-1 is shown in Figure A-2. There are several algorithms for finding the convex hull but for this research, the algorithm by Salim Akl and Godfried Toussaint [Akl78] [Akl79] was chosen.

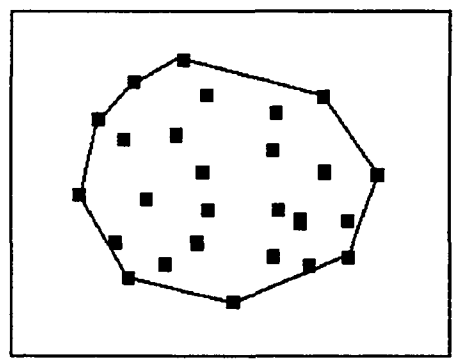


Figure A-2 : The convex hull CH(S) is the region inside the polygon.

Let S be a set of four or more points in a plane which are not on the same line. The steps to determine the convex hull CH(S) are :

Step 1

Identify four points[2] P_1, P_2, P_3 and P_4 (shown in the Figure A-3) from S as indicated below:

P_1 - a point with minimum x coordinate value.

P_2 - a point with maximum y coordinate value.

[2] It is possible that any two consecutive points, of the four points P_1, P_2, P_3 and P_4 may coincide giving us only three distinct points.

P_3 - a point with maximum x coordinate value.

P_4 - a point with minimum y coordinate value.

Each $P_i \in CH(S)$.

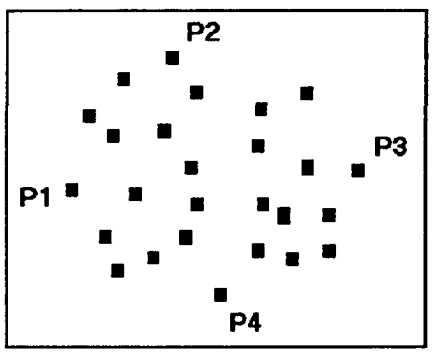


Figure A-3 : Four points at the extremities.

Step 2

Eliminate points of the set S that are inside the quadrilateral $P_1P_2P_3P_4$ from further consideration[3].

Separate the remaining points into four disjoint subsets corresponding to the regions (shown in Figure A-4) defined below:

R_1 : region to left of the directed line segment from P_1 to P_2.

R_2 : region to left of the directed line segment from P_2 to P_3.

[3] The points inside the quadrilateral are not points on the boundary and hence we do not include in our list.

R_3 : region to left the directed line segment from P_3 to P_4.

R_4 : region to left the directed line segment from P_4 to P_1.

A point $P(x,y)$ lies to the left of the directed line segment from $P_1(x,y)$ to $P_2(x,y)$, if $s = (y - y_1)(x_2 - x) + (x_1 - x)(y_2 - y) > 0$.

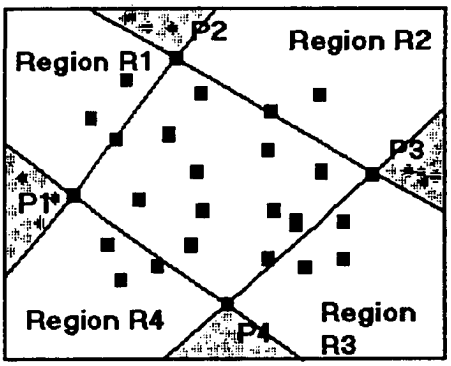


Figure A-4 : Figure shows the four regions. There are no points in the shaded regions

Step 3

Sort the points in regions R_1 and R_2 in ascending order of the x-coordinate and store them in lists L_1 and L_2.

Sort the points in regions R_3 and R_4 in descending order of the x-coordinate and store them in lists L_3 and L_4.

Step 4

For each of the lists L_1 through L_4, eliminate points from the list according to the procedure outlined below:

Consider the first list L_1. Start at the beginning of the list. Consider three consecutive points $P_k(x_k, y_k)$, $P_{k+1}(x_{k+1}, y_{k+1})$, and $P_{k+2}(x_{k+2}, y_{k+2})$, which correspond to three consecutive elements of the list.

Compute $s = (y_{k+1} - y_k)(x_{k+2} - x_{k+1}) + (x_k - x_{k+1})(y_{k+2} - y_{k+1})$.

If $s > 0$ and if the end of the list has been reached stop the procedure otherwise move to the next point, i.e. $k = k + 1$.

If $s \leq 0$, remove $P_{k+1}(x_{k+1}, y_{k+1})$ from the list. If, $P_k(x_k, y_k)$ is not at the beginning of the list, move one backward on the list to the previous point on the list, i.e $k = k - 1$, otherwise move forward to the next point on the list i.e. $k = k + 1$.

Repeat the above step of computing s until the end of the list is reached.

Step 5

Merge the lists L_1, L_2, L_3 and L_4 to generate $CH(S)$.

APPENDIX B

BILINEAR INTERPOLATION

For any image, the gray levels are only defined at integral values of a grid. However after a geometric transformation like rotation[4], shearing etc., the gray level for the transformed image will in general be computed from fractional or non-integral points (coordinates that are not on the grid). Similarly the points at integral coordinates in the original input image may map onto non-integral points on the sampling grid of the output image. This causes image distortions. To smooth any distortions arising from the transformation, gray level interpolation is done. There are several ways gray level interpolation can be done; for this research a bilinear interpolation was chosen. The reason for this choice being that it provides a better interpolation compared to simple linear interpolation and nearest neighbor interpolation. Additionally, the price in terms of programming complexity was small. Compared to other higher order extrapolation methods, it provides a good trade-off in terms of execution time.

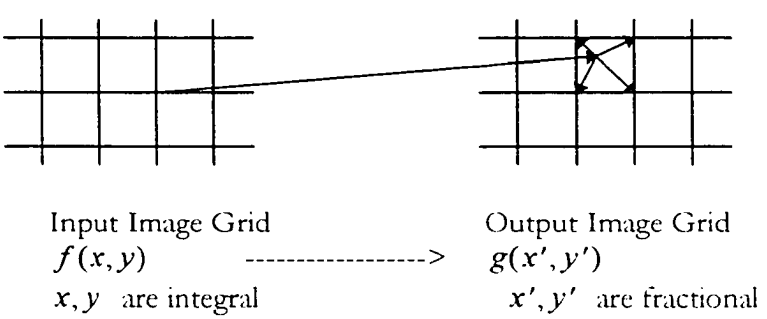


Figure B-1 : Pixel Carryover

[4] All rotation for this research is done using equation 5-1 about the center of the image frame.

There are two possible ways one could transfer the gray levels, using bilinear interpolation. One is the pixel carry-over or forward mapping approach and the other is the pixel filling or backward mapping algorithm. In the forward-mapping approach, when the input pixel maps to a position that is a non-integral coordinate position, its gray level is divided among the four neighboring pixels. This is wasteful because many input pixels map to coordinate positions outside the image frame of the transformed image. In addition each output pixel is addressed several times, with many input pixels contributing to its final gray level. In this research pixel-filling or backward mapping is used. Figure B-1 shows forward mapping and Figure B-2 shows backward mapping.

This Appendix covers the process used for gray level interpolation used in this dissertation. There are two processes:

1. Backward-Mapping Algorithm
2. Bilinear Interpolation

- **Backward-Mapping Algorithm**

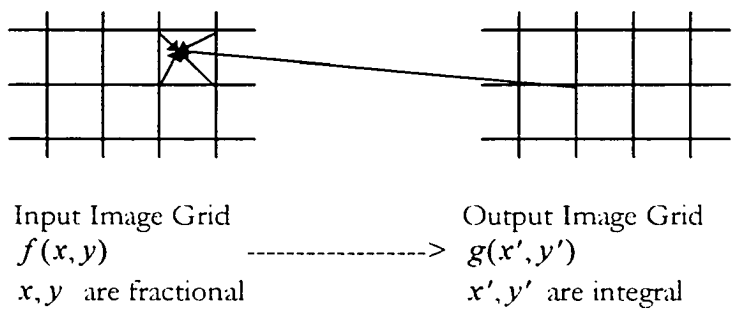


Figure B-2 : Pixel filling

In this approach the output pixels are mapped back into the input image, one at a time, to establish gray levels. If an output pixel falls between four input pixels, its gray level is determined by gray level interpolation, which for the purposes of this research was chosen to be bilinear interpolation. This algorithm generates the output image pixel-by-pixel,

line-by-line. The gray level for each pixel is determined by one interpolation between at most four input pixels.

- **Bilinear Interpolation**

Let the gray level at some point (x', y') on the transformed image map to the gray level at some point (x, y) on the input image. If the point (x, y) is a point that is not an integral coordinate point, then this point is inside a square whose vertices have coordinates that are integral values. Assuming that the point (x', y') on the grid for the output image are integral points, if the corresponding point of the input grid is a non integral point (x, y) we consider the vertices of the square that surrounds this point (x, y). Let us assume that the four vertices are points with coordinates (x_1, y_1), $(x_1 + 1, y_1)$, $(x_1, y_1 + 1)$ and $(x_1 + 1, y_1 + 1)$. We further assume that $f(x, y)$ represents the gray level at any point (x, y) on the input image grid.

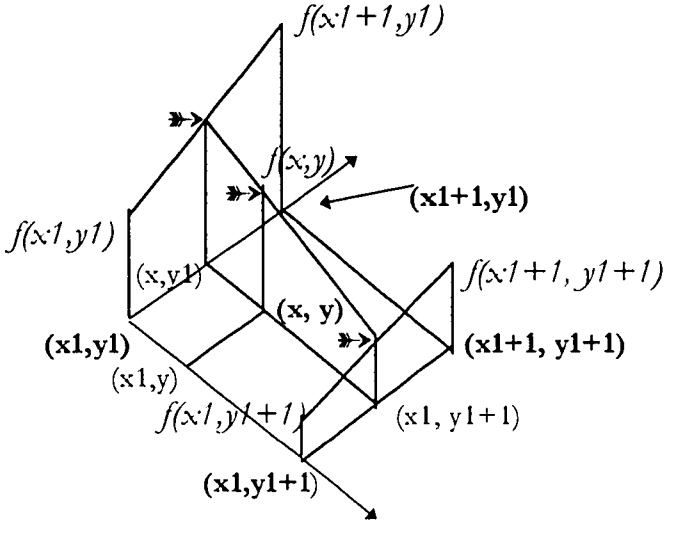


Figure B-3 : Bilinear Interpolation
The four corners of a grid are (x1,y1), (x1+1, y1), (x1,y1+1) and (x1+1, y1+1).

The following set of equations produces a bilinear interpolation function which fits $f(x,y)$ at the corners [Castleman96].

$$f(x,y_1) = f(x_1,y_1) + (x - x_1)[f(x_1 + 1, y_1) - f(x_1, y_1)] \quad \text{Eq. B-1}$$

$$f(x,y_1 + 1) = f(x_1,y_1 + 1) + (x - x_1)[f(x_1 + 1, y_1 + 1) - f(x_1, y_1 + 1)] \quad \text{Eq. B-2}$$

$$f(x,y) = f(x,y_1) + (y - y_1)[f(x,y_1 + 1) - f(x, y_1)] \quad \text{Eq. B-3}$$

Basically, the bilinear interpolation is a hyperbolic paraboloid that intersects all planes parallel to the xz-plane and all planes parallel to the yz-plane in a straight line. It must be noted that when adjacent four corner neighboring pixel gray levels are interpolated using the bilinear algorithm the transformed image has boundaries that are continuous and smoother. Figure B-4 shows the original image of a rectangle that is rotated by 90 degrees, about the center of the image frame using equation Eq. 5-1. Figure B-5 shows the kink in the surface when bilinear interpolation is not used. Figure B-6 shows the same transformed image as in Figure B-1 but with bilinear interpolation.

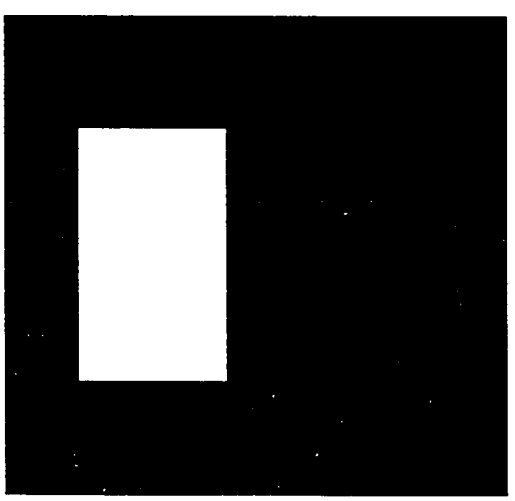


Figure B-4 : Rectangle

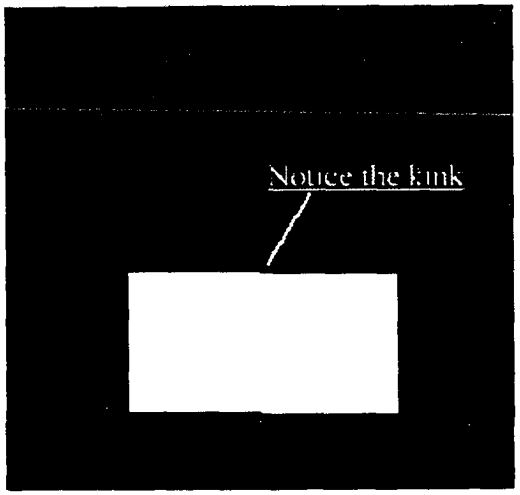


Figure B-5 : Rectangle in Figure B-4 rotated by 90^0. (No interpolation is used)

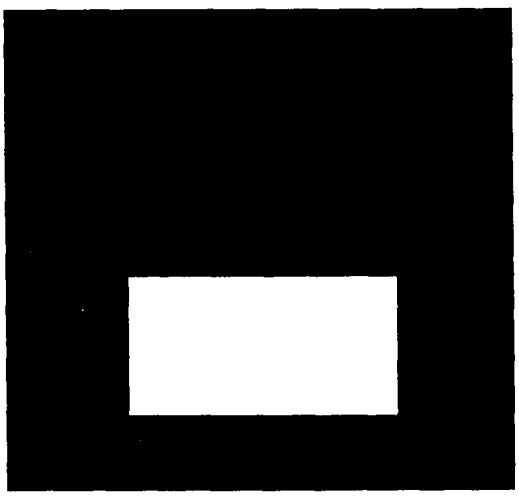


Figure B-6 : Rectangle in Figure B-4 rotated by 90^0 (Bilinear interpolation used).

Table B-1 : The above table shows the pixel gray levels for the image in Figure B-7.

0	0	0	0	0
0	150	150	0	0
0	150	150	0	0
0	150	150	0	0
0	0	0	0	0

Table B-1 shows the gray levels of an image of size 5 by 5 of the rectangle on Figure B-7.

This image on Figure B-7 was rotated by 30^0 about the center of the image frame and is shown in Figure B-8. The gray levels for the 5 by 5 image after rotation is shown in table B-2. Table B-3 shows the number of points in the rotated image, to which each pixel from the original image, contributes gray levels.

Table B-2 : The above table shows the pixel gray levels for the image in Figure B-8.

0	15	40	0	0
0	95	150	60	0
40	150	150	20	0
15	95	75	0	0
0	0	0	0	0

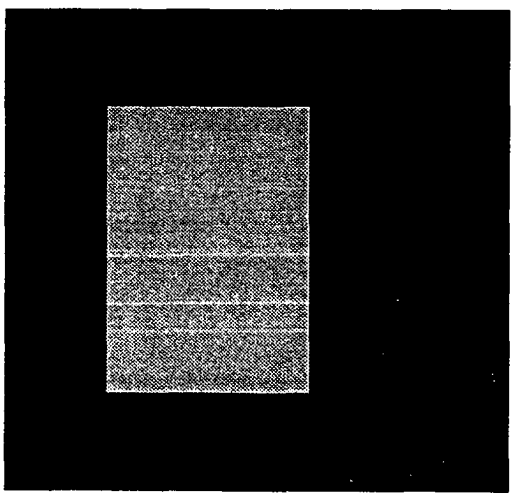


Figure B-7 : Figure shows a rectangle in a 5 in by 5 image.

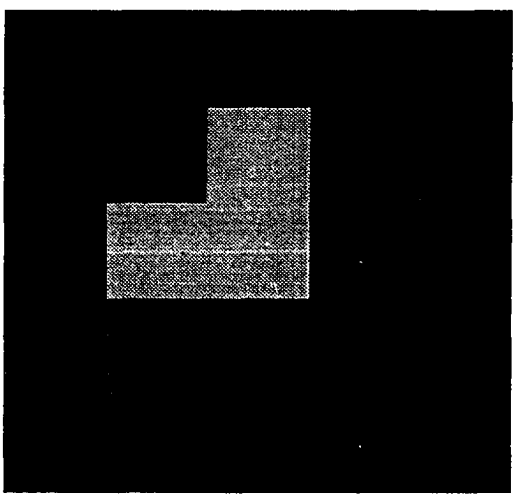


Figure B-8 : Figure shows the rectangle Figure B-7 rotated by 30^0.

Table B-3 clearly indicates that some pixels contribute to more than one pixel in the rotated image. Table B-4 shows the amount contributed by each pixel from the original image to pixels in the rotated image. Table B-3 shows that the contribution of the gray levels is non uniform. Some pixels of the original image contribute to as many as 5 pixels

in the rotated image. Additionally some pixels contribute a lot more gray level value than other pixels as shown in table B-4.

Table B-3 : The $(i,j)^{th}$ entry in the table shows the number of pixels of the rotated image, (shown in Figure B-8), to which a contribution is made, by the $(i,j)^{th}$ pixel of the original unrotated image (shown in Figure B-7).

1	2	2	2	1
2	4	4	4	2
2	4	5	5	2
2	4	5	5	2
1	2	2	2	1

Table B-4 : The $(i,j)^{th}$ entry in the table shows the amount of gray level contribution made to rotated image, (shown in Figure B-8), by the $(i,j)^{th}$ pixel of the original unrotated image (shown in Figure B-7).

0	0	0	0	0
0	155.384740	135.288599	0	0
0	135.288599	190.192343	0	0
0	155.384740	135.288599	0	0
0	0	0	0	0

This essentially results in a redistribution of gray levels. This redistribution causes distortion of the objects in the image. However for objects with high resolution, this distortion is not observable under visual inspection. This is demonstrated with the help of Figures B-9 and B-10. Figure B-9 shows the image of a rectangle which has a much higher resolution. Figure B-10 shows the same image rotated by 30^0. The rotated figure does not show the distortions and gray level distributions, when the image are visually inspected.

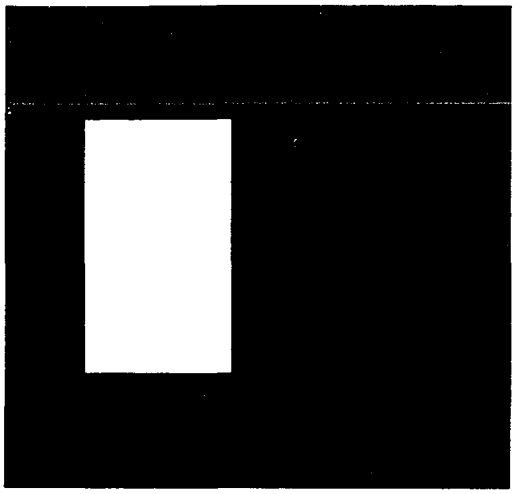


Figure B-9 : Figure shows the image of a rectangle

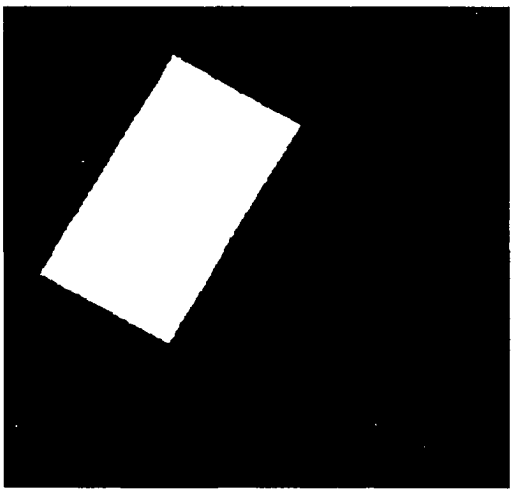


Figure B-10 : Figure shows the image of the rectangle in Figure B-9 rotated by 30^{0} about the center of the image frame

APPENDIX C

PRINCIPAL COMPONENT ANALYSIS

C.1 Introduction

Principal Component Analysis (PCA) is concerned with explaining the variance-covariance structure through linear combinations of the original variables. Its general objectives are:

- data reduction and
- interpretation.

To examine the relationship among a set of p correlated variables, it may be useful to transform the original set of variables into a new set of uncorrelated variables called principal components. If the original variables are linearly dependent, it will result in fewer uncorrelated variables, resulting in the reduction of the dimensionality of the problem. These new variables are derived in decreasing order of importance. Principal Component Analysis is essentially a variance minimization technique. The first principal component accounts for as much of the variation as possible in the original data, second principal component accounts for as much of the remaining variation as possible, and so on. The transformation is orthogonal.

If the original variables are nearly uncorrelated, then there is no point in carrying out a PCA for dimensionality reduction. In the case of a two dimensional image, the inherent assumption in using Principal Component Analysis is that the image data is at the least somewhat spatially ellipsoidal. In this case the principal axis of the ellipse would account for most of the spatial variance. Principal Component Analysis would identify the axis as

the principal components. Figure C-1 illustrates the ellipse fitting the data points and the principal components.

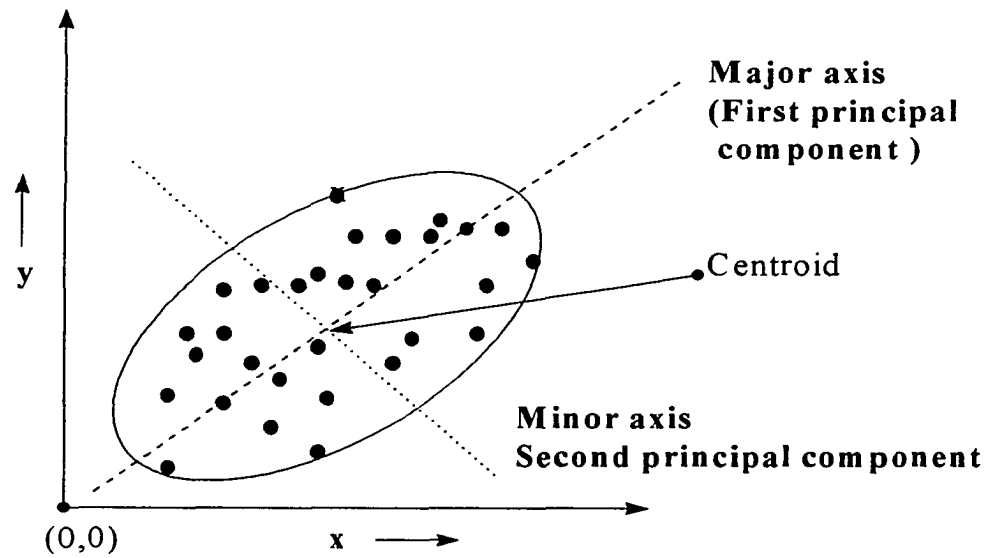


Figure C-1 : The dotted line shows the principal components for the set of points shows as dots in the figure.

Algebraically, principal components $Y_1, Y_2, Y_3, \ldots Y_p$ are linear combinations of the p random variables $X^T = [X_1, X_2, \ldots X_p]$ with mean μ and covariance matrix Σ.

$$Y_1 = a_{11}X_1 + a_{12}X_2 + \ldots\ldots + a_{1p}X_p \qquad \text{Eq. C-1}$$

$$Y_2 = a_{21}X_1 + a_{22}X_2 + \ldots\ldots + a_{2p}X_p \qquad \text{Eq. C-2}$$

.....

$$Y_p = a_{p1}X_1 + a_{p2}X_2 + \ldots\ldots + a_{pp}X_p \qquad \text{Eq. C-3}$$

$$\mathrm{Y_j} = \mathbf{a}_\mathrm{j}^\mathrm{T}\mathrm{X} \qquad \text{Eq. C-4}$$

where $\mathbf{a}_j^T = [a_{j1}, a_{j2}, a_{j3}, \ldots a_{jp}]$ is a vector of constants. To eliminate the indeterminacy due to scale, the coefficient vector $\mathbf{a}$ is restricted to unit length. This is done by imposing the orthogonality condition

$$\mathbf{a}_j^T \mathbf{a}_j = 1. \qquad \text{Eq. C-5}$$

This particular normalization procedure ensures that the distances in the $p-spaces$ (the original domain of the variables as contrasted to the transformed domain) are preserved [Chatfield80].

The first principal component Y_1 is the linear combination with maximum variance, determined by choosing $\mathbf{a}_1$ and hence the linear combination $\mathbf{a}_1^T \mathbf{X}$ that maximizes

$$Var(Y_1) = Var(a_1^T X) = E[a_1^T (X-\mu)(X-\mu)^T a_1] = a_1^T E[(X-\mu)(X-\mu)^T] a_1 = a_1^T \Sigma a$$

Eq. C-6

subject to $\mathbf{a}_1^T \mathbf{a}_1 = 1$.

The second principal component is the linear combination $\mathbf{a}_2^T \mathbf{X}$ that maximizes

$$Var(Y_2) = Var(a_2^T X) = E[a_2^T (X-\mu)(X-\mu)^T a_2] = a_2^T E[(X-\mu)(X-\mu)^T] a_2 = a_2^T \Sigma a_2$$

Eq. C-7

subject to the following conditions :

1. $\mathbf{a}_2^T \mathbf{a}_2 = 1$ and Eq. C-8

2. $Cov(Y_1, Y_2) = Cov(a_1^T X, a_2^T X) = 0$. Eq. C-9

The i^{th} principal component is the linear combination $\mathbf{a}_i^T \mathbf{X}$ that maximizes

$$Var(Y_i) = Var(a_i^T X) = E[a_i^T (X-\mu)(X-\mu)^T a_i] = a_i^T E[(X-\mu)(X-\mu)^T] a_i = a_i^T \Sigma a_i$$

Eq. C-10

subject to the following conditions :

1. $\mathbf{a}_i^T \mathbf{a}_i = 1$ and Eq. C-11

2. $Cov(Y_i, Y_k) = Cov(a_i^T X, a_k^T X) = 0$ for all $k < i$. Eq. C-12

To maximize the variance function subject to one or more constraints, one can use the method of Lagrange multipliers[Edwards94]. We formulate the Langrangian function $L(a)$ as shown in the equation

$$L(a_1) = a_1^T \Sigma a_1 - (a_1^T a_1 - 1)\lambda \quad \text{Eq. C-13}$$

Differentiating, with respect to a_1, we get

$$\frac{\partial L}{\partial a_1} = 2 \Sigma a_1 - 2 a_1 \lambda \quad \text{Eq. C-14}$$

Setting the above equation to zero, we have

$$(\Sigma - I\lambda) a_1 = 0 \quad \text{Eq. C-15}$$

To have a solution for a_1 that is not a null vector, $(\Sigma - I\lambda)$ must be singular. λ must be chosen so that

$$|\Sigma - I\lambda| = 0 \quad \text{Eq. C-16}$$

The covariance matrix Σ is a square symmetric matrix and positive semidefinite[5]. The eigenvalues of the covariance matrix Σ are real and nonnegative.

Let the covariance matrix Σ have the eigenvalue-eigenvector pairs (λ_1, e_1), (λ_2, e_2), ... $(\lambda_p e_p)$ where $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \lambda_p \geq 0$. The i^{th} principal component Y_i satisfies the following conditions:

$$Var(Y_i) = e_i^T \Sigma e_i = \lambda_i \text{ where } i = 1,2,3..,p \quad \text{Eq. C-17}$$

$$Cov(Y_i, Y_k) = e_i^T \Sigma e_k = 0 \text{ for } i \neq k \quad \text{Eq. C-18}$$

---

[5] A symmetric matrix Σ and the quadratic form $a^T \Sigma a \geq 0$ for all a.

The principal components are uncorrelated and have variances equal to the eigenvalues of Σ.

Total population variance =

$$\sigma_{11} + \sigma_{22} + \ldots \sigma_{pp} = \sum_{i=1}^{p} Var(X_i) = \lambda_1 + \lambda_2 + \ldots + \lambda_p = tr(\Sigma) = \sum_{i=1}^{p} Var(Y_i)$$

Eq. C-19

where $\sigma_{ii} = Var(X_i)$.

The proportion of total variance due to the k^{th} principal component is equal to

$$\frac{\lambda_k}{\lambda_1 + \lambda_2 + \lambda_3 + \ldots + \lambda_p}.$$

A more detailed proof of the above results is found in most multivariate statistics text books [Chatfield80][Cooley71] and hence is not replicated here. This research uses these results in determining the principal components of the modulus maxima image.

C.2 Computation

This part of the appendix deals with the steps to compute the first principal component and the angle relative to the $x-axis$ for a set of points in a plane.

Let the gray level for each pixel in any image be represented by $f(x_i, y_i)$ where x_i and y_i are the pixel coordinates in the horizontal and vertical direction. Steps to compute the principal components for a two dimensional binary image are :

- **Step 1**

  Compute the centroid (μ_x, μ_y) of the object in the image.

  The centroid (μ_x, μ_y) is computed using the equations given below. We only consider points for which $f(x_i, y_i) \neq 0$.

$$\mu_x = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{Eq. C-20}$$

$$\mu_y = \frac{1}{n}\sum_{i=1}^{n} y_i \quad \text{and} \qquad \text{Eq. C-21}$$

n is the number of pixels for which $f(x_i, y_i) \neq 0$.

- **Step 2**

Compute the covariance matrix $\Sigma = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{bmatrix}$. We only consider points for which $f(x_i, y_i) \neq 0$.

$$\sigma_{xx} = \frac{1}{n}\sum_{i}^{n} x_i^2 - M_x^2 \qquad \text{Eq. C-22}$$

$$\sigma_{yy} = \frac{1}{n}\sum_{i}^{n} y_i^2 - M_y^2 \qquad \text{Eq. C-23}$$

$$\sigma_{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - M_x M_y \quad \text{and} \qquad \text{Eq. C-24}$$

n is the number of pixels for which $f(x_i, y_i) \neq 0$.

- **Step 3**

Compute the largest eigenvalue λ_1. This can be done using the equation

$$\lambda_1 = \frac{(\sigma_{xx} + \sigma_{yy}) + \sqrt{(\sigma_{xx} + \sigma_{yy})^2 - 4(\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2)}}{2} \qquad \text{Eq. C-25}$$

- **Step 4**

Compute the corresponding eigenvector for the eigenvalue λ_1 using the equation given below:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ (\lambda_1 - \sigma_{xx})/\sigma_{xy} \end{bmatrix}. \qquad \text{Eq. C-26}$$

This eigenvector is the first principal component. It is the major axis of the object. The minor axis is orthogonal to this axis in the same plane.

- **Step 5**

  Compute the angle of the major axis relative to the $x-axis$. This angle is given by the equation shown below:

$$\theta = a\tan\left[\frac{\lambda_1 - \sigma_{xx}}{\sigma_{xy}}\right] \qquad \text{Eq. C-27}$$

C.3 Usage

In this dissertation we compute the principal components of the two objects being registered. The difference in the angle of the first principal component with the $x-axis$ is computed.

APPENDIX D

INTRODUCTION TO EDGE DETECTION

D.1 Introduction

An edge is the boundary between two relatively distinct gray level regions. The underlying assumption made by edge detection techniques is that each of the regions separated by the edge are sufficiently homogeneous that the transition between two regions can be determined on the basis of only gray level discontinuities. Most edge detection techniques determine edges by computing an approximate local derivative operator.

Figure D-1 shows the gray level profile, the first and second derivatives of a function that represents the gray level profiles. The first derivative of the gray level profile is positive at the leading edge of a transition, negative at the trailing edge, and zero in areas of constant gray level. The second derivative is positive for the transition part associated with the dark side of the edge, negative for the transition part associated with the light side of the edge, and zero in areas of constant gray level. Hence, the magnitude of the first derivative can be used to detect the presence of an edge, and the sign of the second derivative can be used to determine whether an edge pixel lies on the dark side or light side. The second derivative changes sign (a zero crossing) at the midpoint of a transition in gray level and this information provides an approach for locating edges in an image.

If an $n \times n$ image $f(x,y)$ is treated as consisting of n rows and n columns we can find the first and second derivatives in the horizontal and vertical directions at each pixel location (x,y) in the image. The gradient of the image $f(x,y)$ at location (x,y) is given by

$$\nabla f = \begin{bmatrix} \partial f(x,y)/\partial x \\ \partial f(x,y)/\partial y \end{bmatrix}. \qquad \text{Eq. D-1}$$

The gradient vector points in the direction of maximum rate of change of $f(x,y)$ at (x,y).

The direction of the gradient vector represents the direction of the edge and is given by the expression :

$$\text{Edge direction} = \tan^{-1}\left(\frac{\partial f}{\partial y} \Big/ \frac{\partial f}{\partial x} \right). \qquad \text{Eq. D-2}$$

For an image with digitized pixels, the difference between adjacent pixel values is used to approximate the derivative. The different commonly used edge detectors differ in how they compute partial derivatives and the gradient at every pixel location. The computation of the edges is accomplished by spatial convolution of the digital image data using a kernel or set of masks.

Using a weighted average of each pixel in the input image and its eight surrounding neighbors a value is computed for the output image at this pixel location. This group of pixels used to compute a weighted average is called the kernel. The weights used to compute the weighted average represented as a 3×3 array is called the convolution mask. Each edge detector makes use of a different 3×3 convolution mask. Spatial convolution with these masks essentially results in high pass spatial filtering. The high pass filtering yields convolution coefficients that are used enhance high frequency components and to remove low frequency components. This procedure results in edge enhancement and edge detection.

Consider the 3×3 mask M show below:

$$M = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & m_8 \end{bmatrix}$$

Let $f(x,y)$ represent the gray level at pixel location (x,y) for an $n \times n$ input image. Let $O(x,y)$ represent the same gray value at location (x,y) in the output image after convolution of this $n \times n$ input image using this 3×3 the mask. The output gray values $O(x,y)$ for every pixel location (x,y) in the output image, is computed as:

$$O(x,y) = \sum_{i=0}^{2} \sum_{j=0}^{2} f(x-1+j, y-1+i) m_{(j+3i)} \qquad \text{Eq. D-3}$$

Gray levels in the output image for pixels that are on the image frame boundary are computed from only those pixels that are inside the image frame of the input image.

D.2 Mask Types

D.2.1 Laplacian Edge Detector

The Laplacian edge detector is a scalar operator for $2-D$ functions. It is defined as:

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \qquad \text{Eq. D-4}$$

The Laplacian edge detector finds all edges regardless of their direction. The Laplacian operator generates sharper peaks at edges than many of the edge detecotrs discussed in this appendix. It is implemented using any one of the two kernels shown below:

Mask # 1

$$M = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Mask # 2

$$M = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

D.2.2 Sobel Edge Detector

The Sobel edge detector consists of two masks. One mask provides better information about horizontal edges and the other mask provides better information about vertical edges. Every input image pixel is convolved using both masks. The maximum value obtained after both convolutions are done, is used as the pixel gray level in the output image for that pixel location. The result is called an edge magnitude image.

Mask # 1

$$M = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Mask # 2

$$M = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

D.2.3 Prewitt Edge Detector

The Prewitt Edge detector also uses two masks. Every pixel in the input image is convolved with both masks. The maximum value obtained after the two convolutions is used as the pixel gray level in the output image for that pixel.

Mask # 1

$$M = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Mask # 2

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

D.2.4 Kirsch Edge Detector

The Kirsch Edge detector consists of eight masks. Each pixel in the input image is convolved using all eight masks. The maximum pixel value obtained after the eight convolutions is the pixel gray level in the output image for that pixel.

Mask # 1

$$M = \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

Mask # 2

$$M = \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}$$

Mask # 3

$$M = \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}$$

Mask # 4

$$M = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$$

Mask # 5

$$M = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}$$

Mask # 6

$$M = \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}$$

Mask # 7

$$M = \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix}$$

Mask # 8

$$M = \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

D.2.5 Roberts Edge Detector

The Roberts edge detector also uses a difference technique but uses a different spatial convolution than was used by the other edge detectors in this Appendix. It uses two distinct differences to represent a finite approximation to the derivative of brightness. The direction of the edges is computed in the diagonal directions and the magnitude is combined by taking the square root of the sum of the squares of the magnitude.

Let $f(x,y)$ represent the gray level at pixel location (x,y) in an $n \times n$ input image. Let $O(x,y)$ represent the gray value at location (x,y) in the output image after convolution of the input image. The output $O(x,y)$ for every pixel location (x,y) in the output image for the Roberts Edge Detector, is computed as shown:

$$O(x,y) = \sqrt{(f(x,y+1) - f(x+1,y))^2 + (f(x+1,y+1) - f(x,y))^2} \ . \qquad \text{Eq. D-5}$$

D.3 Observation

The edge detectors presented in this Appendix are commonly used edge detectors. All these edge detectors magnify noise [Russ94][Baxes94][Gonzalez92].

VITA

Raj Sharman was born in Bombay, India. He received his bachelor's and master's degrees from the Indian Institute of Technology, Bombay, India. He received a second master's degree from Louisiana State University, United States of America. He became a citizen of the United States of America in 1992. During the course of doing his doctoral studies he found the Lord Jesus Christ. He became a follower of the Lord Jesus Christ and accepted Christianity as a way of life. His interests include Image Processing, Computer Vision, Automated Detection and Classification, Multi Sensor Data Fusion, Database systems, Pattern Recognition, Information Theory and Information Systems, Electronic Commerce and finding solutions to Scientific and Engineering problems. He will receive the degree of Doctor of Philosophy in August 1998.

DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Raj Sharman

Major Field: Computer Science

Title of Dissertation: Wavelet Based Registration of Medical Images

Approved:

John M. Tyler
Major Professor and Chairman

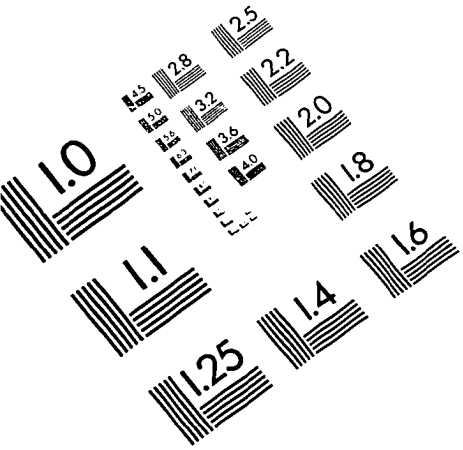
Dean of the Graduate School

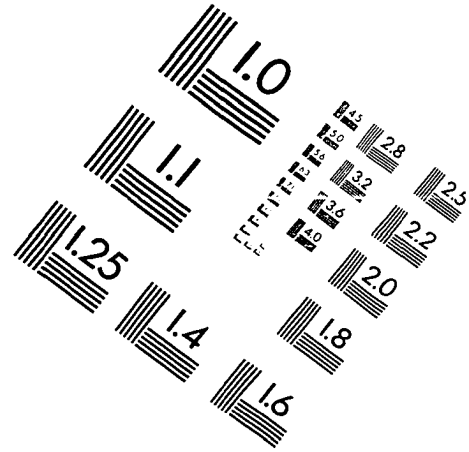
EXAMINING COMMITTEE:

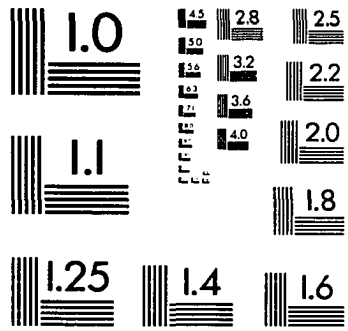
Date of Examination:

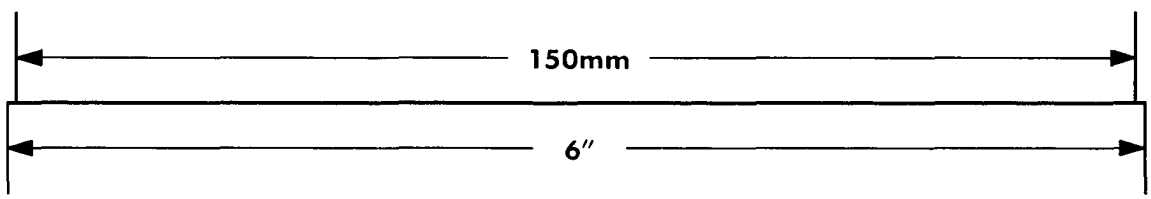
May 8, 1998

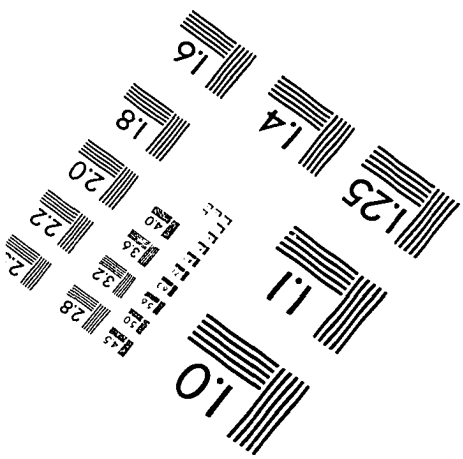
IMAGE EVALUATION
TEST TARGET (QA-3)











APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989