Landmark detection in MR brain images using SURF

Jayachandran, G.

Award date:
2012

Technische Universiteit
Eindhoven
University of Technology

MASTER THESIS

Landmark detection in MR brain images using SURF

Author
Ganesh Jayachandran
(0755553)
M.sc Embedded Systems

g.jayachandran@student.tue.nl

Supervisor from Philips Research
Dr. Ahmet Ekin
ahmet.ekin@philips.com

Supervisor from TU/e
Prof. Dr. Ir. Gerard De Haan
G.d.Haan@tue.nl

August 1, 2012

Abstract

Practitioners in the area of neurology often use brain landmarks in critical applications ranging from planning stereotactic neurosurgery to the study of disease progression over time. Although manual identification of these landmarks by experts is often possible for one or more datasets, the task becomes impractical for larger datasets where the manual annotation gets more erroneous, time-consuming, and expensive. As a result, developing fully automatic methods to localize brain landmarks has been an active area of research in the field of medical image analysis. Among the existing work, a group of algorithms considers solving the landmark detection problem by using an image registration algorithm. Although this approach is applicable to detection of different landmarks, depending on the registration method and the landmark of interest, it can be too computational or may lack accuracy. In contrast, a second group of algorithms aims to develop a specialized detection algorithm for each landmark. These algorithms achieve better accuracy and improved speed in the detection of the landmark they are designed for. However, their extension to the detection of another landmark often requires a new algorithm. In this thesis, we propose a framework that bridges these two approaches and strives for a good level of generalizability and accuracy in the detection of brain landmarks. The proposed framework is composed of two stages: In the analysis stage, we first represent landmarks by local feature points and identify reference landmarks (defined as those that can be reliably detected) in a number of annotated brain MR datasets. This stage can be considered as the learning stage of our framework and is done once for a defined landmark. In the second stage, we consider landmark detection as a special form of feature point matching guided by learned reference landmarks. In this stage, in addition to the particular landmark, we make use of the more reliably-detectable reference landmarks in the neighborhood of the landmark-of-interest. The proposed framework is generic in the sense that we can apply the same scheme for any landmark of interest. It has also improved accuracy and consistent detection performance because of the use of reference landmarks. In this work, we used SURF (Speeded Up Robust Feature) feature extraction and matching technique to represent and identify landmarks; however, any local feature point extraction method is also immediately usable in the developed scheme. The proposed framework has been tested on both T1 and T2 weighted MR images and for different applications, such as slice matching, volume matching, and mid-sagittal plane detection. The experiments show a significant improvement in the landmark detection with the proposed scheme over the detection without reference landmarks for both T1- and T2-weighted MR images.

Keywords: *SURF, landmark detection, search and retrieval, volume matching, mid sagittal plane detection*

Contents

Chapter 1

Introduction

1.1 Problem description

Anatomical landmarks are well-defined references in the anatomy that experts use to establish biologically meaningful correspondences between structures [1]. The human brain consist of such structures that are important in the field of neuroimaging. Identification of anatomical landmarks are used in critical application such as planning stereotactic and functional neurosurgery [2], for the localization of neuroanatomic targets, structure segmentation and labeling in neuroradiology [3], to study brain asymmetry caused by pathology [4] and to study specific disease progression over time.

Although manual identification of these landmarks by experts is often possible for one or more datasets, the task becomes impractical for larger datasets where the manual annotation gets more erroneous, time-consuming, and expensive. As a result, developing fully automatic methods to localize brain landmarks has been an active area of research in the field of medical image analysis. Among the existing work, a group of algorithms considers solving the landmark detection problem by using an image registration algorithm. Although this approach is applicable to detection of different landmarks, depending on the registration method and the landmark of interest, it can be too computational or may lack accuracy. In contrast, a second group of algorithms aims to develop a specialized detection algorithm for each landmark. These algorithms achieve better accuracy and improved speed in the detection of the landmark they are designed for. However, their extension to the detection of another landmark often requires a new algorithm.

1.2 Solution approach

The approach to address the above problem is to develop a generic framework where we first analyze brain MR images to identify the most reliable landmarks and use the identified landmark positions to detect other landmarks. The framework will employ computer vision techniques to extract information from MR brain images. The extracted information will be used to identify the desired landmarks.

1.3 Contributions

The important contributions towards this work include:

- Developing a generic framework by

  (a) Analyzing brain MR volumes and extracting the most reliably detectable reference landmarks of the brain by using robust features.

  (b) Using the reference landmarks to detect a landmark-of-interest in brain MR images.

- Applying feature extraction algorithm from computer vision to MR brain images.
- Contrast independent detection – applicable for T_1 and T_2 images

Chapter 2

Brain landmarks and feature extraction techniques

In this chapter an introduction to brain landmarks and the existing methods to detect brain landmarks are presented. The localization of clinically important landmarks in brain images is crucial for many neurological studies. In Section 2.1 brain landmarks are discussed. In Section 2.2 a brief insight on Magnetic Resonance Imaging (MRI) and the inherent problems with MRI are explained. In Section 2.3 the existing methods to detect brain landmarks are discussed. Later in Section 2.4, the feature extractors that were considered for this project are explained.

2.1 Brain landmarks

The human brain is a relatively small organ (around 1400g) sitting within the skull and protected by membranes called the meninges, which include an external dense outer layer, called the dura mater, a thin inner layer, called the pia mater, and an intermediate layer, the arachnoid, constituted as a layer of fibers. The brain floats in the cerebrospinal fluid (CSF) that has various functions such as protection, nourishment and draining [5].

The basic subdivisions of brain are the two cerebral hemispheres, the brain stem, and the cerebellum [5] as shown in Figure 2.1. The junction between the top of the spinal cord and the brain is called the brain stem and it comprises around ten percent of the central nervous system. This part is essential for involuntary functions such as breathing, digestion, heart rate and blood pressure. The cerebellum (little brain) is posteriorly connected to the brain stem. Despite occupying only one tenth of the volume of brain, the cerebellum accounts for around half of the total number of neurons. Its primary functions are movement that it contributes to coordination, precision, and accurate timing of motor activity and balance. Receiving and processing a range of inputs from the eyes, ears, balancing systems and cortex, the cerebellum dispatches instructions back through the brain stem to other regions of the brain.

Found above the midbrain and between the large cerebral hemispheres is the diencephalon region, which contains several important substructures, including the thalamus and the hypothalamus. The egg-shaped thalamus is essential in the gating, processing and transfer of almost all the sensory information (except from the nose) entering the brain [6]. Sitting be-

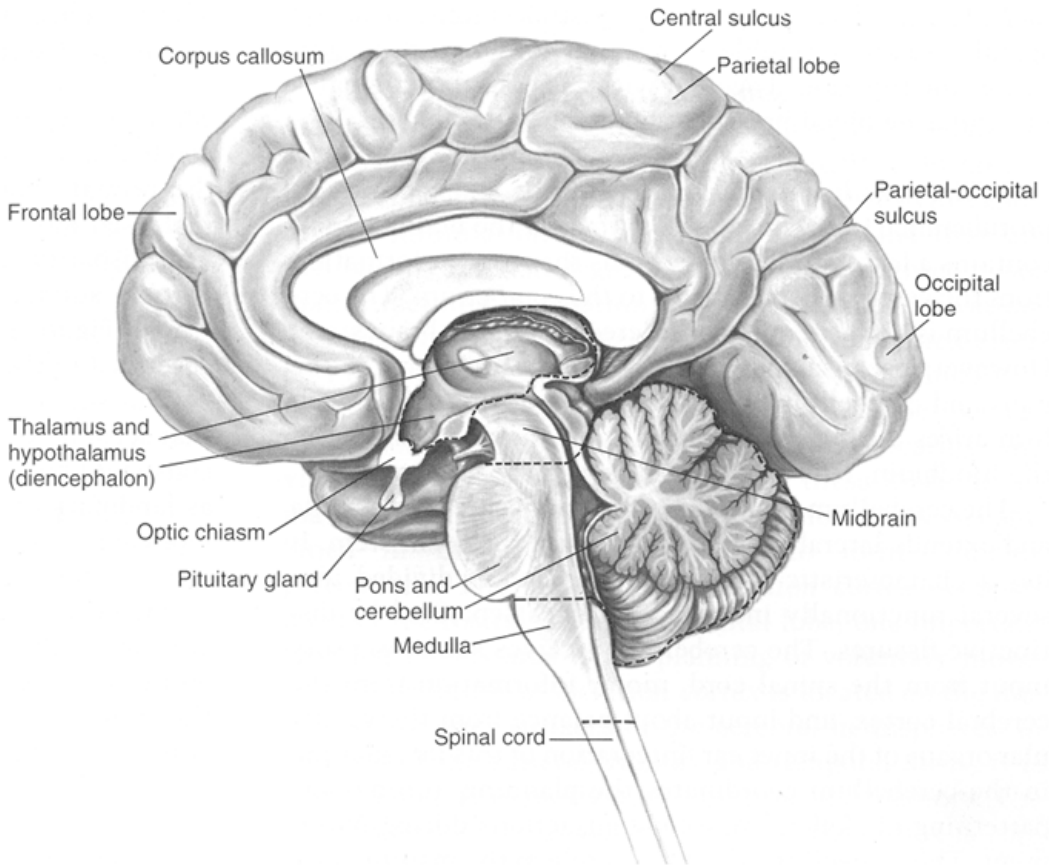


Figure 2.1: Cross Section – Human Brain [5]

neath the thalamus is the hypothalamus, where a multitude of critical functions are controlled and regulated. Connected to almost every other part of the brain, the hypothalamus is essential to motivation, including the seeking out of activities that the person finds rewarding, like sex and music or even drugs. As a regulator of hormone release, the hypothalamus is also involved in everything from homeostasis and eating to maternal behavior. It also manages the daily cycle of the body: the circadian rhythm [6].

The two cerebral hemispheres are separated from one another by a longitudinal fissure, also called the interhemispheric fissure or the mid sagittal plane, which contains the falx cerebri, a membranous septum that separates the hemispheres [5] as shown in Figure 2.2. The cerebral hemispheres are symmetrical, and their surfaces, called the cortex or gray matter, contain the cellular bodies of the neurons. The surfaces of the hemispheres are highly convoluted and can be described as a succession of crests, called the gyri,and fissures separating them, called the sulci [5]. Underlying this gray mantle is the white matter, which is made of bundles of fibers emerging from the bodies of neurons. These fibers are the axons enveloped by their myelin. The two hemispheres are connected by a broad commissure of white-matter tracts: the corpus callosum [5].

The hemispheres are commonly subdivided into six lobes, four of which are shown in Fig-

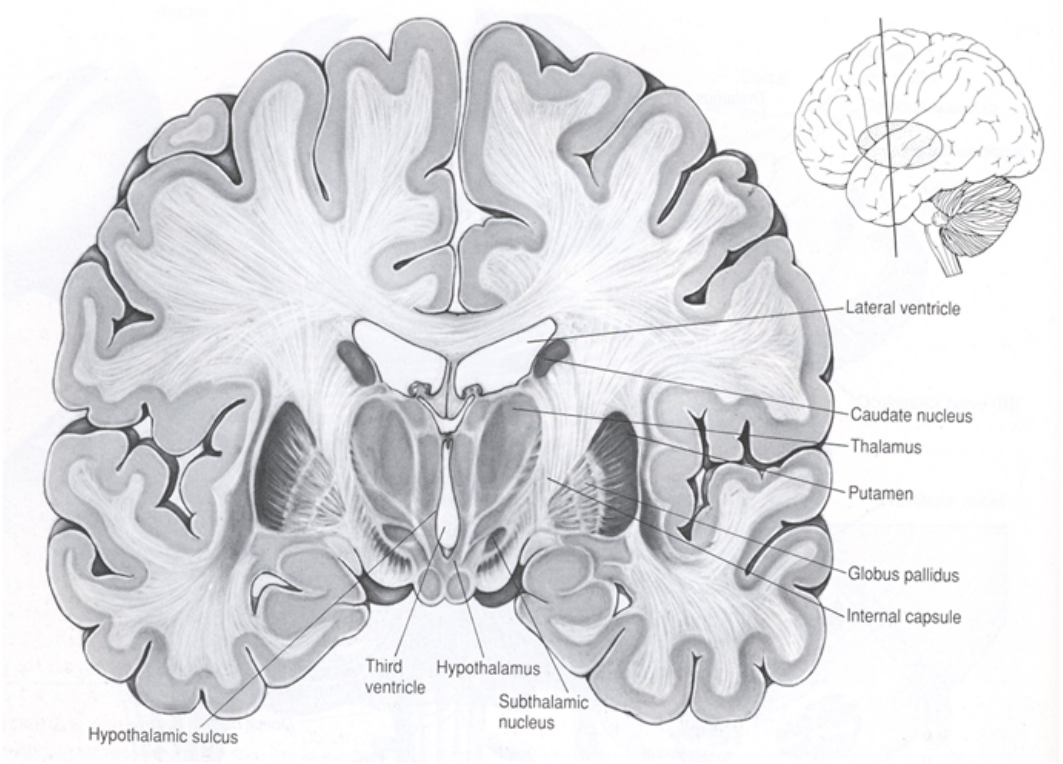


Figure 2.2: Cross Section – Cerebrum

ure 2.3, that were named after the bones of the skull overlying them [8]. The frontal lobe, directly beneath the forehead, is involved in what are collectively termed higher functions: attention, planning, language and movement. It is like a master control unit that helps to integrate information and govern what the rest of the brain does. Behind it, at the top of the head, the parietal lobes process lots of sensory information, allowing us to perceive the world and our place within it [6]. At the back, the occipital lobes deal primarily with vision; it is here that signals from the eyes become transformed into a useful visual representation. Finally, the temporal lobes down each side of the brain focus on sound and language, and, by way of their connection with the hippocampus, memory formation and retrieval.

The hemispheres of brain are held together by the corpus callosum (thick body). This is the largest bundle of nerve fibres in the body and the main channel through which information flows from one side of the brain to the other - as it must for the whole brain to function properly [6].

Landmarks refer to any brain structure that can be represented as a 3D structure like the cerebellum, corpus callosum, etc., a plane such as the mid sagittal plane or a point like the anterior commissure (AC) and posterior commissure (PC).

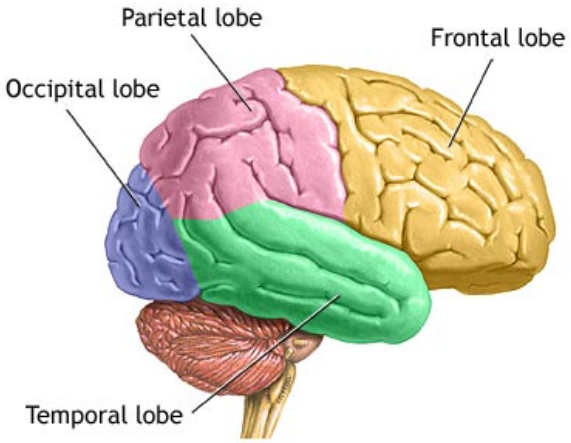


Figure 2.3: Brain Lobes [7]

2.2 Challenges in detecting landmarks in MRI images

Magnetic resonance imaging (MRI) is the preferred imaging modality to non-invasively take pictures of brain, by its cross section. MRI has several advantages:

- Neither the constant magnetic field, nor the radio frequency modulated one, which are present in the MRI device, represent a danger to the humans.
- The MRI signal is determined by the proton density of the tissue, T_1 and T_2 relaxation times, the type of sequence used, and the selected acquisition parameters. These parameters give the opportunity to enhance the image contrast between the two tissues by cleverly choosing the type of sequence and the acquisition parameters, and thus optimize the differentiation between tissue structures.
- Some of the MRI devices give us the opportunity to produce multi–spectral images, that is two or more images of the same cross section with different parameter setting.

In order to prevent possible misinterpretations, most MR images are acquired such a way that the tissue contrast of various images is determined mainly on a single tissue parameter. In this context, T_1,T_2 and PD–weighted images are produced. The MRI acquired is a volumetric image that contain voxel forming the basic volumetric element, slices that is composed of several volume elements or voxels and volume that is composed of several slices. All of these spatial units can be a landmark of interest. The slice landmark is particularly of interest to this work, since a slice landmark is well represented by its local content. This can be used to localize a structure to a particular region of the brain volume or with in a slice landmark. The MR images are stacked as slices in axial, coronal and sagittal direction as shown in Figure 2.4. The axial plane is an x-y plane, parallel to the ground that separates the superior from the inferior. A coronal plane, is an x-z plane, perpendicular to the ground that separates the anterior from the posterior. A sagittal plane, is an y-z plane, perpendicular to the ground that separates left from right. Figure 2.5(a) and 2.5(b) shows the three basic anatomical planes and the brain structures observed in the planes. Figure 2.6 depicts the T_1

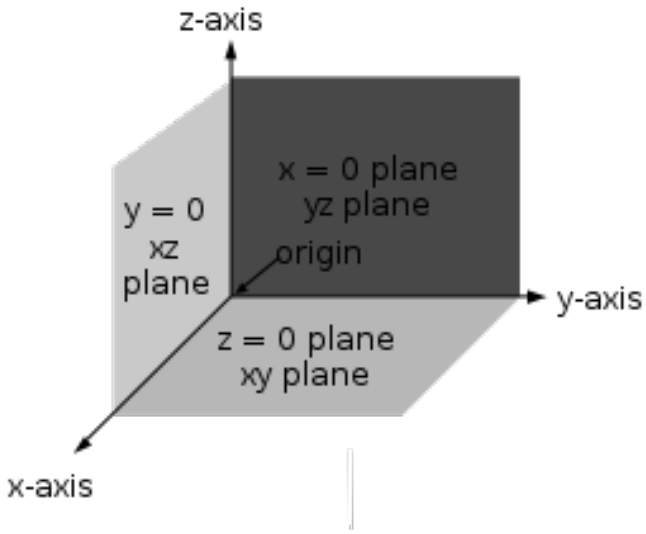
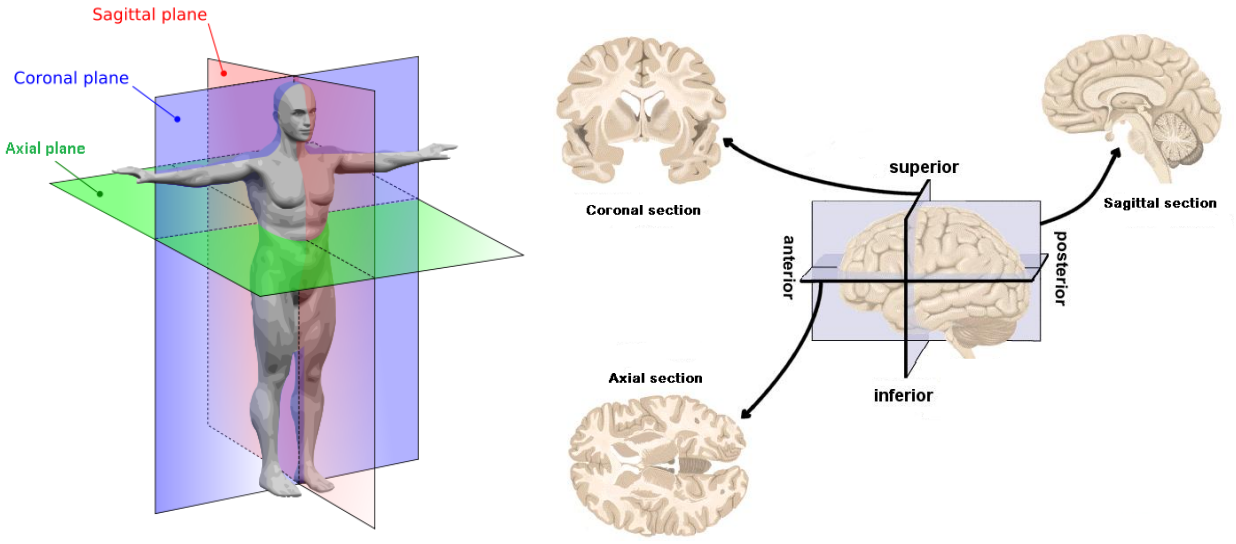


Figure 2.4: MRI volume showing axial(xy plane), sagittal(xz plane) and coronal(yz plane) direction [9]

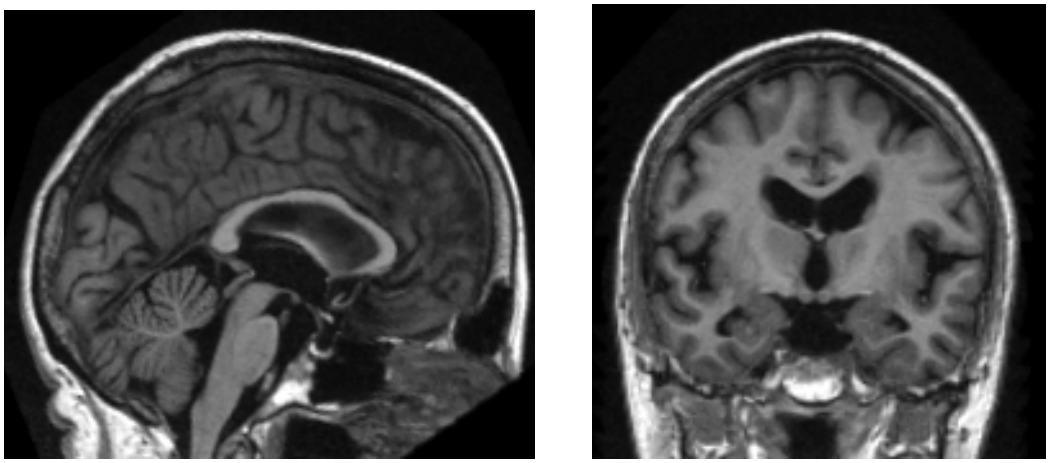
weighted images of brain.



(a) (b)

Figure 2.5: MRI planes. (a) Three basic anatomical planes [10]. (b) Typical cross section of the brain

Magnetic resonance images (MRIs) acquired with similar protocols but on different scanners show dissimilar intensity values for the same tissue types [11]. At times the recording protocol also vary leading to differences in scale and voxel spacing from one patient to another. The former limits the use of intensity information while the later reduces the accuracy while comparing different MR volumes. The rotation induced in the acquired MRI due to



(a) (b)

Figure 2.6: MRI. (a) MRI cross section of human brain. (b) MRI cross section of cerebrum.

pose difference from patient to patient while recording the MRI make it difficult for matching algorithms.

Low frequency noises in the magnetic resonance imaging are generally referred to as intensity inhomogeneity or intensity non–uniformity(INU) artifact, and manifest as a smooth intensity variation across the image. This phenomenon causes regions of the image belonging to the same tissue class have different intensities. Low magnitude INU is hardly noticeable for the human eye, but it can induce confusion for image processing algorithms with high sensitivity.

The *partial volume artifact*(PVA), also known as partial volume effect(PVE), represents a phenomenon that is present in MR medical images due to its coarse resolution. Even if MRI reportedly has higher than other medical imaging techniques, the highest resolution MR images still have only 1–3 pixels per millimeter. Under such circumstances, it is unavoidable to have pixels that are shared between two or even among three or more tissue types. This is a problem for segmentation algorithms and may influence the extraction accuracy of some features.

High similarity between relevant and irrelevant segments in brain MR volumes, such as brain tissue vs. non-brain structures in the same slice and relevant slice vs. neighboring slices in the same volume, makes it difficult for detecting landmark location. Anatomical differences and pathological changes between patients, make it difficult for matching landmarks [12].

2.3 Existing methods to detect landmarks in brain

Quantitative and qualitative studies of anatomical brain tissues and structures that have distinctive structural or functional properties usually relies on accurate detection of brain landmarks. Software tools such as MRIcro [13], ITK-SNAP [14], rview [15] and 3D-Slicer [16] are some of the freely available tools to perform manual segmentation of brain landmarks.

The current proprietary software tools for segmentation of brain landmarks from MRI scans are very laborious and time consuming and requires highly trained experts/technicians. An experienced researcher may require two hours to trace a single structure such as the hippocampus, and more than a week to trace all of the major structures of the brain [17]. Differences in criteria among experts can lead to systematically different volume estimates of some brain regions, and so the highest consistency and sensitivity is achieved when a single individual traces the entire dataset. However, the criteria used by even a trained expert can subtly drift during the course of a long study. For these reasons, semi automatic and fully automated procedures for identifying brain landmarks have attracted considerable interest [17].

Morey et al. [17] has shown that automated procedures identify brain structures accurately than manual procedures. Atlas-based registration is a commonly used technique to segment 3D or 2D landmark from a MRI volume. In principle, an atlas consists of two components: the gray-scale information and the segmentation containing labels for different tissue classes or anatomical structures. After registering an atlas to the destination image volume using gray-scale information, the resulting transformation is used to propagate the labels into the space of the destination volume [18]. Many different approaches has been reported using atlas-based registration for segmenting anatomical structures structures [18, 19, 20]. The major drawback of the atlas based landmark detection is that, it is computationally expensive and the results depend on the proper selection of atlases.

Unay et al. [12] presented a novel search and retrieval technique for finding relevant slice landmarks in brain MR (magnetic resonance) volumes. The approach uses Local Binary Pattern(LBP) [21] and KLT Feature Points [22] to retrieve the slice landmarks. But the technique is applied only in 2–D and relies on mid sagittal plane detection before matching.

Mid-sagittal plane (MSP), a plane landmark that is generally used to study brain symmetry. The existing methods for MSP detection follow either a feature based or a symmetry-based approach. Ekin [23] proposed a robust feature-based approach to make it possible to detect the mid-sagittal plane as long as two image lines are not affected by pathalogical abnormality. In the symmetry-based approach, mid-sagittal plane is defined as the one that maximizes the similarity between the brain and its reflection. The methods following this approach [24, 25, 26] first define a parameter space that describes the MSP, a similarity measure, such as cross-correlation, to assess the interhemispheric symmetry in the selected feature space, such as intensity or edges [25], and a search method and search criteria to find the parameters that maximize the similarity measure.

Kruggel and Yves [27] have proposed an image processing chain to align MR brain datasets with stereotactical coordinate system [28]. The method combines brain peeling, detection of AC and PC and normalization in the Talairach space. The AC and PC point landmarks in particular are detected by spanning a constant gradient field between the brain hemispheres and identifying the local flow peaks in the mid-sagittal plane. Hu et al. [29] proposed to identify modified Talairach cortical landmarks [30] in three steps. In the first step, three planes containing the landmarks are calculated. Then the plane is thresholded and segmented with the chosen thresholds and morphologic operations. Finally, the segmentation is refined to compensate for the influence of optic nerves and the partial volume effect [29].Prakash et al. [31] have proposed a two step method to identify the AC and PC point landmark widely

used for image registration and in neurosurgery. The method uses the anatomy information surrounding the AC and PC to identify the third ventricle. In the first step 2D data is processed to obtain the initial position of the commissures and the second step refines the initial positions by processing 3D volume of interest. Fu et al. [32] proposed automatic identification of fourth ventricular landmarks which are used to describe a variability study of brain stem structures. The method uses the technique proposed by Prakash et al. [31] to identify the AC and PC and then use these landmarks to identify the fourth ventricular landmarks.

The method proposed by Prakash et al. [31] and Fu et al. [32] rely on a successful localization of corpus callosum. To overcome this problem Ardekani and Bachman [33] proposed a model based approach to detect AC and PC. The approach uses 3D templates derived from multiple model images for training. The training can be extended to identify other landmarks which is close to the proposal made in this report [33]. Recently Don et al. [34] have employed artificial neural networks to detect AC and PC which has shown to minimize the AC-PC localization error rate. The existing methods to detect point landmark have a major drawback that they work only for that particular landmark. To detect a different point landmark, a new method has to be designed. In this thesis we aim to develop an extendable or a generic method that can be used to detect all landmarks.

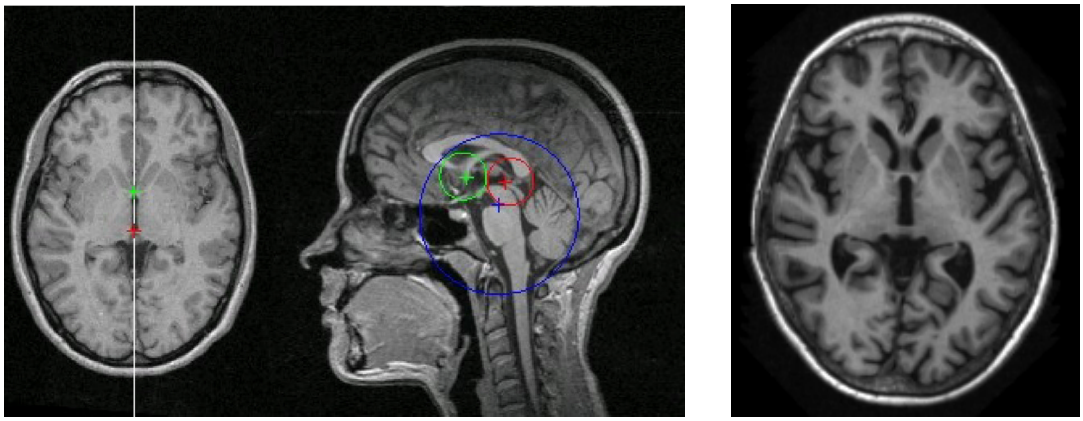
2.4 Existing feature extraction techniques

In MRI brain volumes the landmarks are well distinguished from its neighbors based on their intensity values. Hence the MRI brain landmarks can be identified across volumes by their intensity values that form the feature set for this project. MR brain images are subjected to various challenges as explained in Section 2.2. So feature extractors that are robust to noise, bias field, rotation and that is generalizable to any dataset can be employed. Figure 2.7 shows a point and a slice landmark. It is observed that point landmark is well distinguished from its neighbors in Figure 2.7(a).

The feature extractors that are robust and have repeatable features were considered for this project. These feature extractors perform well in case of rotation up to 10 degrees [35]. Also the detected feature set depend on the local information rather than the global information, hence the extractors are robust. These feature extractors are widely used in computer vision applications but are rarely used in the field of medical image processing particularly in MR images. In the case of medical image processing these feature extractors are applied in image registration [36][37] and in image search and retrieval [38].

2.4.1 Scale Invariant Feature Transform (SIFT)

Scale Invariant Feature Transform(SIFT) was first proposed by Lowe [39]. SIFT describes image features that have many properties that make them suitable for different images of an object or a scene. The features are invariant to image scaling and rotation and partially invariant to illumination and camera 3D viewpoint. They are well localized in both the spatial and frequency domains, reducing the probability of disruption by occlusion, clutter, or noise [39]. Since typically an image consist of a large number of features, extracting these would be costly. SIFT reduces the cost of feature extraction by taking a cascade filtering approach,



(a) (b)

Figure 2.7: MRI brain landmark. (a) Point landmark (AC-PC) [33]. (b) Slice landmark.

in which expensive operations are applied only at locations that pass an initial test [39]. The following are the major stages stated by Lowe [39] that are used to compute features:

- Scale-space extrema detection: The first stage of computation searches over all scales and image locations. It is implemented efficiently by using a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation.
- Keypoint localization: At each candidate location, a detailed model is fit to determine location and scale. Keypoints are selected based on measures of their stability.
- Orientation assignment: One or more orientations are assigned to each keypoint location based on local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations.
- Keypoint descriptor: The local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

Various refinements on this basic scheme have been proposed. Ke and Sukthankar [40] applied PCA on the gradient image. This PCA-SIFT yields a 36-dimensional descriptor which is fast for matching [40], but proved to be less distinctive than SIFT in a comparative study by Mikolajczyk and Schmid [41] and slower feature computation reduces the effect of fast matching. In the same paper the authors have proposed gradient location and orientation histogram (GLOH) which is even more distinctive with the same number of dimensions. However, GLOH is computationally more expensive.

2.4.2 Features from Accelerated Segment Test (FAST)

FAST(Features from accelerated segment test) is a feature detector based on segment test proposed by Rosten and Drummond [42]. This is sufficiently fast that it allows on-line operation for tracking features. A test is performed for a feature at a pixel p by examining a

Table 2.1: Time taken for feature detection on a PAL field (768 × 288) pixels on the test system. [42]

Detector	FAST	Harris Corner
Time (ms)	2.6	44

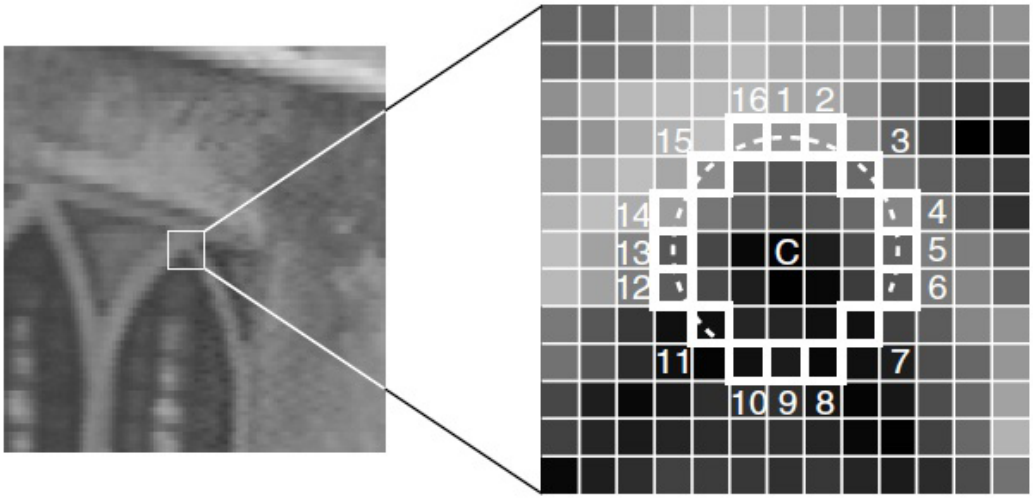


Figure 2.8: FAST Feature detection in an image patch. The highlighted squares are the pixels used in the feature detection. The pixel at C is the centre of a detected corner. [42]

circle of 16 pixels (a Bresenham circle of radius 3) surrounding p. A feature is detected at p if the intensities of at least 12 contiguous pixels are all above or all below the intensity of p by some threshold, t [42]. This is shown in Figure 2.8.

The test for this condition is optimized by examining pixels 1, 9, 5 and 13, to reject candidate pixels more quickly, since a feature can only exist if three of these test points are all above or below the intensity of p by the threshold. With this optimization, on a sample sequence of video, the algorithm examines, on average, 3.8 pixels to test if there is a feature at a given location [42]. Table 2.1 shows the time taken to compute the corner by Harris corner detector and FAST on a PAL field on the test system. From the Table 2.1 it is observed that FAST is approximately 17 times quicker than Harris corner detection. FAST is not scale or rotation invariant compared to SIFT and SURF.

2.4.3 Speeded Up Robust Features (SURF)

Speeded Up Robust Features (SURF) was proposed by Bay et al. [43]. SURF approximates or even outperforms previously proposed schemes with respect to repeatability, distinctiveness, and robustness, yet is computed and compared much faster. The following are the major stages in SURF as stated by Bay et al. [43] :

- Interest point detection: The interest points are selected at distinctive locations in

the image, such as corners, blobs, and T-junctions. The most valuable property of an interest point detector is its repeatability. The SURF detector uses Hessain matrix but uses a very basic approximation which are found to be more stable and repeatable compared to the Harris based detectors. Also, approximations like the Difference of Gaussians(DoG) [44] brings speed at a low cost in terms of lost accuracy.

- Feature description: The neighborhood of every interest point is represented by a feature vector. SURF descriptor uses Haar-wavelet responses within the interest point neighborhood. This descriptor is distinctive and, at the same time, robust to noise, detection errors, and geometric and photometric deformations.

Both the detector and descriptor uses integral images to reduce the computation time. Comparison between SURF and SIFT on various parameter is shown in Table 2.2. The results in the table are based on the work of Juan and Gwun [35] who performed the experiments on various parameters with large dataset.

Table 2.2: Comparison between SURF and SIFT on various parameters [35]

Method	Time	Scale	Rotation	Blur	Illumination	Affine
SURF	Best	Good	Common	Good	Best	Good
SIFT	Common	Best	Best	Best	Common	Good

Chapter 3

Proposed Framework

In this Chapter we present the proposed framework that can be used to detect any brain landmark. The landmark can be a voxel, a slice or a group of slices(volume). The proposed framework model is shown in Figure 3.1. The framework has two stages (a) Analysis stage and (b) Application stage. In the analysis stage, we first represent landmarks by local feature points and identify reference landmarks (defined as those that can be reliably detected) in a number of annotated brain MR datasets. This stage can be considered as the learning stage of our framework and is done once for a defined landmark. In the second stage, we consider landmark detection as a special form of feature point matching guided by learned reference landmarks. In this stage, in addition to the particular landmark, we make use of the more reliably-detectable reference landmarks in the neighborhood of the landmark-of-interest. The following sections brief about the analysis and application stages of the framework.

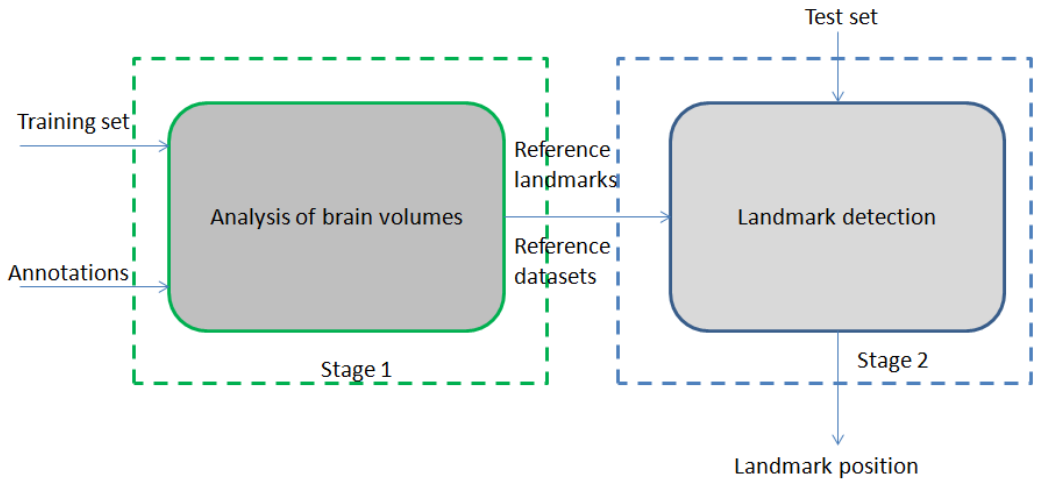


Figure 3.1: Proposed Framework

3.1 Analysis stage of the framework

In the analysis stage, the brain MR volumes are analyzed to identify reference landmarks and reference datasets. The Figure 3.2 shows the flow chart of analysis stage. The input to

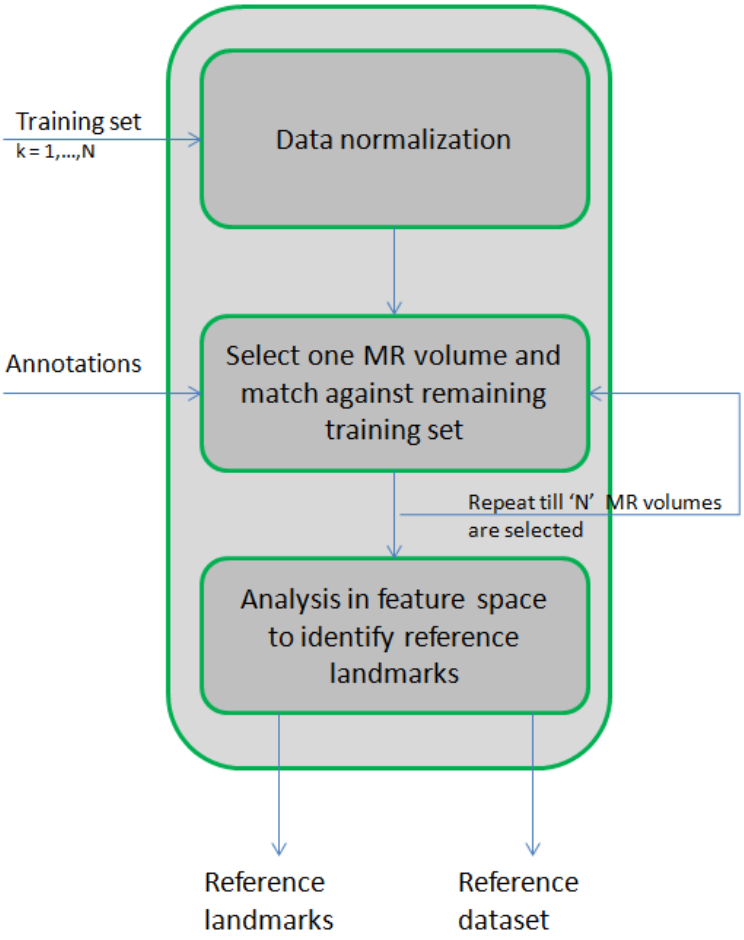
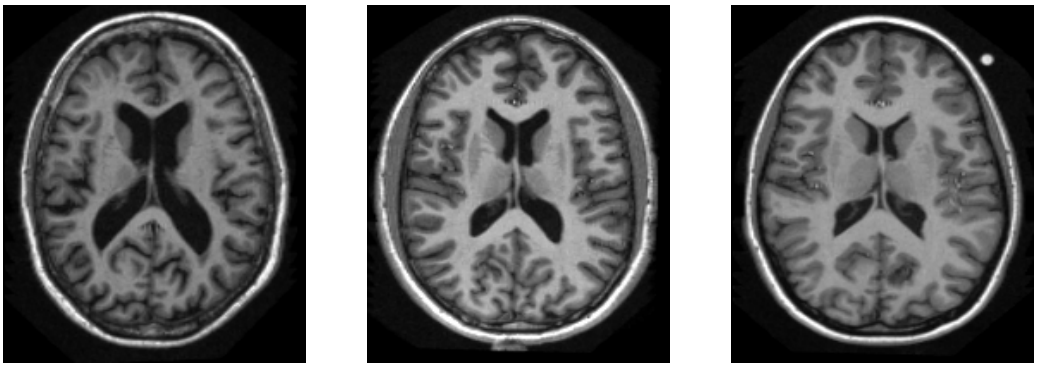


Figure 3.2: Flow chart of the analysis stage.

the framework is the annotated data and the training set. The annotations can be manually provided by experts or can be automated by image registration algorithms. Although manual annotation of these landmarks by experts is often possible for one or more datasets, the task becomes impractical for larger datasets where the manual annotation gets more erroneous, time-consuming, and expensive. Alternatively the accuracy of automatic annotation depends on the registration technique and landmark of interest. The number of MR volume in training set in this project is limited to 11, but in principle there can be any number of MR volumes in the training set. The training set is normalized to accommodate difference in MRI settings while recording the MRI scan. Each of the MR volume in the training set act as a reference at-least once. The training set is analyzed in the feature domain by matching the local feature points of the reference and the training set. In principle the features can be from SIFT, SURF or any feature extractor that extracts robust and repeatable features. In this project we use SURF features. The result of such an analysis is a set of reference landmarks positions and reference datasets. We select the reference datasets, because the reference landmarks are not only defined by their positions but they are also dependent on the features themselves. Hence we select reference datasets such that they contain all features of reference landmarks. The number of reference datasets can be two or three or four or N depending on the reference

landmarks. The reference datasets are chosen such that they do not represent similar brain sizes. Figure 3.3 shows an example of reference datasets around the ventricle region of brain showing different sizes of brain.



(a) (b) (c)

Figure 3.3: (a-c)Slices around the ventricle region from three datasets. They represent large, medium and small ventricles of brain.

3.2 Application stage of the framework

In the application stage the framework uses the knowledge gained in the analysis stage to reliably identify the query landmarks in a test MR volume. The Figure 3.4 shows the flow chart of the application stage. In addition to the application specific input(voxel/slice/group of slices), we make use of the more reliably-detectable reference landmarks in the neighborhood of the landmark-of-interest identified in the analysis stage. The framework first matches the reference landmarks positions from the reference dataset in the test MR volume, identifying the reference landmark(s) positions in the test MR volume. The framework then estimates the queried landmark position(application specific input) using the identified reference landmarks in the test MR volume. This gives an initial position estimate of the query landmark and an initial search area within which the queried landmark is contained. The final position of the queried landmark is found by localizing the initial position estimate within the initial search area.

In this Chapter, the reference landmarks and reference datasets are shown as two separate outputs of analysis stage for clarity purpose. From now on the reference landmarks and the reference datasets will be referred together as reference landmarks, and reference landmark selection inherently means the reference datasets are also selected representing different brain sizes as explained in the previous section.

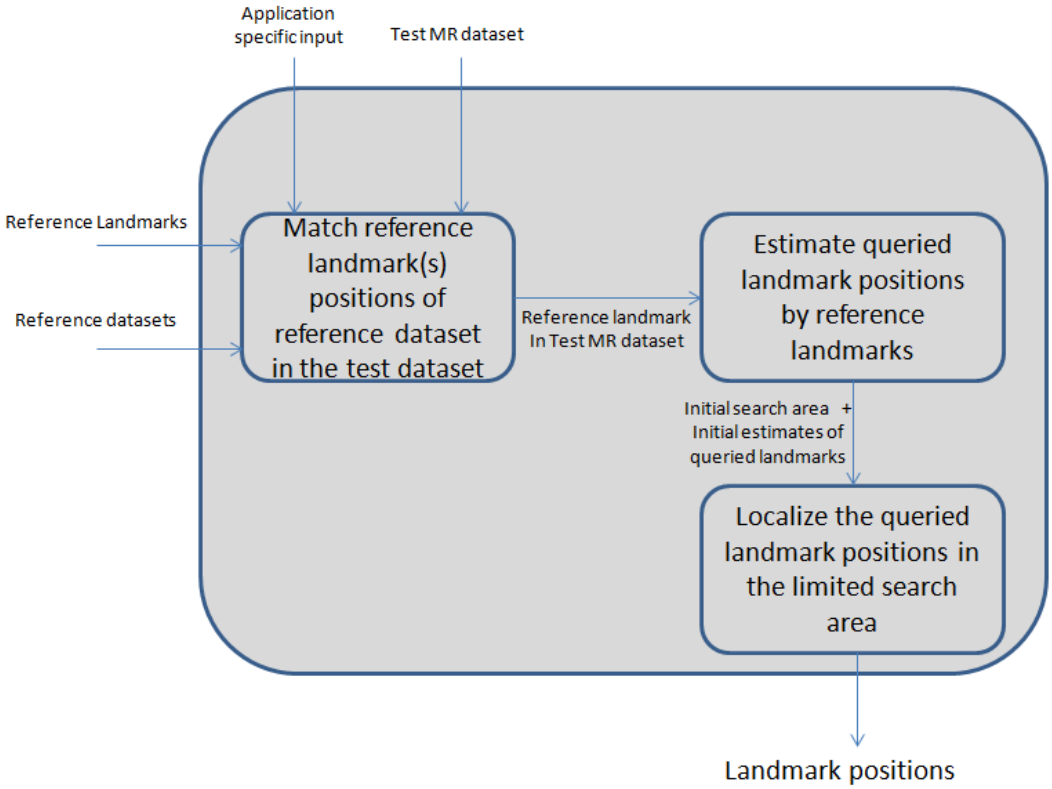


Figure 3.4: Flow chart of the landmark detection stage.

Chapter 4

Normalization and initial landmarks detection

Medical images are often deteriorated by noise due to various sources of interference and other phenomena that affect the measurement processes in imaging and data acquisition systems. The nature of the physiological system under investigation and the procedures used in imaging also diminish the contrast and the visibility of details. In MR brain images, the MRI settings vary according to the clinical practice. This leads to difference in the dynamic range of the acquired scan and also the voxel dimension. To handle these differences spatial and intensity normalization techniques are employed in this project. In Section 4.1, a brief explanation on spatial normalization is presented. In Section 4.2, intensity normalization and its application in this project are discussed. In Section 4.3, we identify the initial landmarks that form the MRI boundaries.

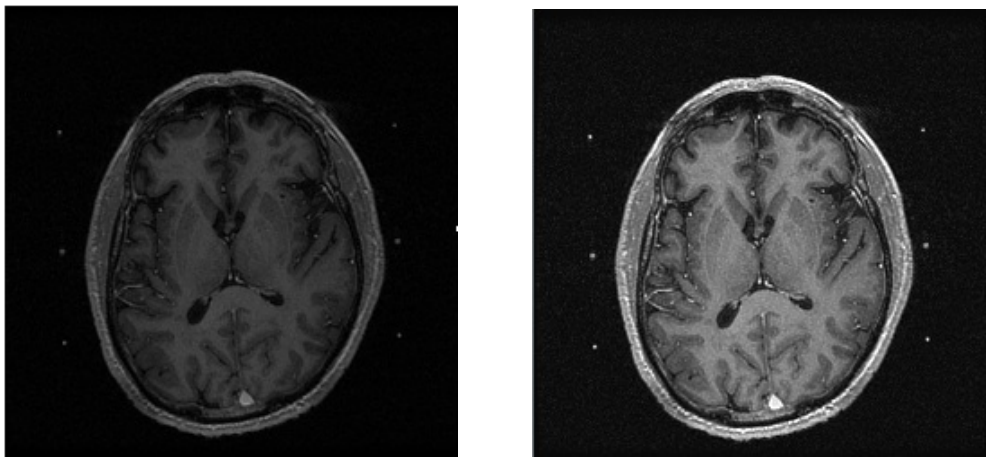
4.1 Spatial normalization by voxel size correction

MRI settings vary according to different clinical practice. Hence, there may be wide variations in the recording parameters. One such parameter is the dimension of the voxels that can influence feature extraction. For instance an MR volume acquired from one scanner can have voxel dimension of 1.0×1.0×1.0 mm and an MR volume from different scanner may have a size of 0.94×0.8×1.0 mm. This can reduce the accuracy of feature extraction and matching. To avoid this, the voxel dimensions are normalized to 1.0×1.0×1.0 mm. Thus all MR volumes are handled in the same manner in the rest of the framework. The normalization is performed by linear interpolation as it produces acceptable results in a speedy manner.

4.2 Intensity normalization

MRI acquired from various sources, as well as from the same source but at different time points, generally do not have similar intensity ranges. For instance, MRI acquired from one scanner may have a dynamic range of 0–100 whereas the MRI acquired from another scanner may have a dynamic range of 0–3500. To handle such non-standard intensity variations, intensity normalization is performed. The normalization is done in two steps:

1. Background estimation.



(a) (b)

Figure 4.1: Dynamic range. (a) Axial slice with dynamic range 0–150. (b) Normalized image with range 0–255.

2. Re-scaling pixels of interest to a fixed range, such as 0–255.

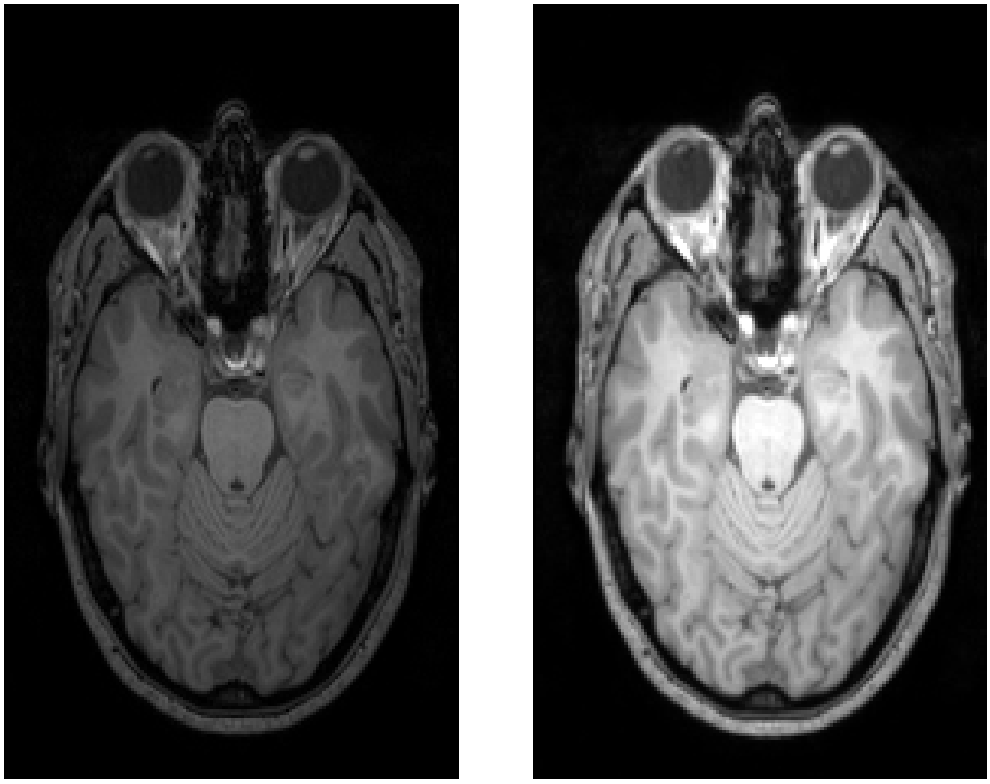
4.2.1 Background estimation

The background for an MR scan is defined as the air surrounding the patient while recording the MRI. The background estimation is based on the histogram of the MRI volume. Instead of calculating the histogram for the complete volume, the estimation is done at the slice level i.e. local histogram estimation. Figure 4.1(a) and 4.2(a) shows the dynamic range of an axial slice taken from different MRI sources. Intuitively, it can be observed that structures of interest(the brighter pixels) are well distinguished from the background(the dark pixels). The histogram corresponding to the axial slice in Figure 4.1(a) is shown in the Figure 4.3(a). It is observed the pixels representing the dark pixels or the background is contained with in 6% of the histogram value.

4.2.2 Pixel intensity re-scaling

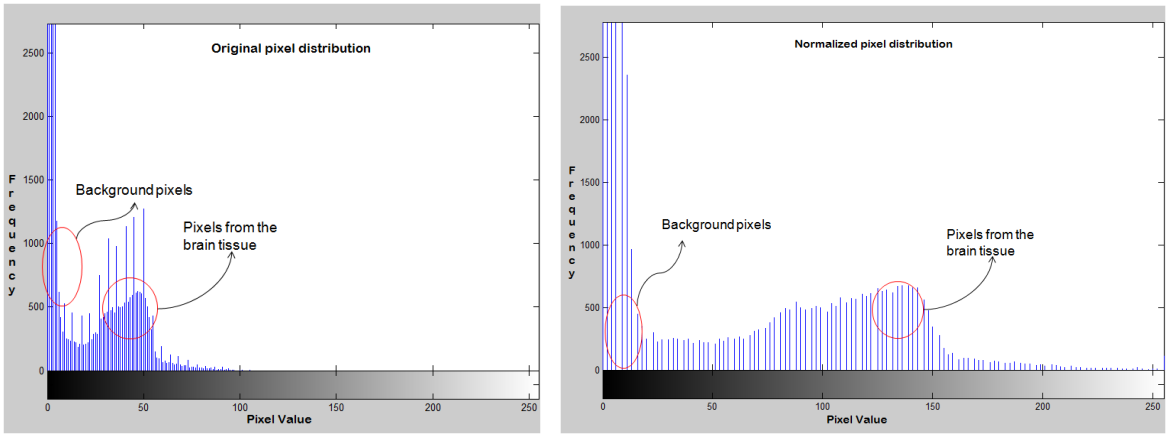
Once the background pixels are known, a window is defined to scale the pixel value within the window. The window lower limit is set to 8% of the histogram value thus eliminating the pixels in the background. The upper limit is selected such that the brightest pixels contributing to the skull is excluded. This is observed at around 80% of the histogram value. The pixels that lie within this window are re-scaled to 0–255. The pixels below the lower window limit are clipped to 0, while the pixels that fall outside upper window limit are clipped to 255. The scaling operation for a pixel "x" is defined by the following equation:

$$Rescale(x) = \left\{ \begin{array}{ll} Outputmin & \text{x} < \text{windowmin}, \\ Outputmax & \text{x} > \text{windowmax}, \\ Outputmin + (x - Windowmin) \times \frac{Outputmax - Outputmin}{Windowmax - Windowmin} & \text{otherwise} \end{array} \right\} \tag{4.1}$$



(a) (b)

Figure 4.2: Dynamic range. (a) Axial slice with dynamic range 0–927. (b) Normalized image with range 0–255.



(a) (b)

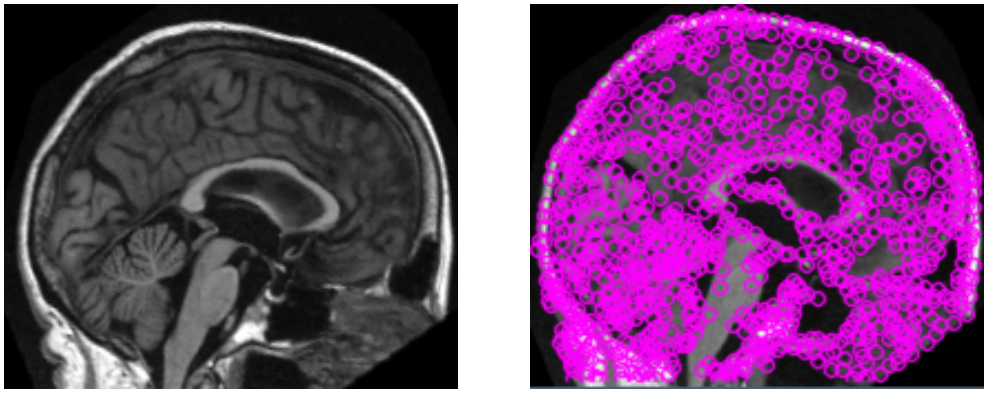
Figure 4.3: Histogram of unprocessed and normalized images. (a) Histogram of an axial slice with full range. (b) Histogram of an intensity normalized axial slice.

The window minimum and window maximum are set dynamically based on the analysis of the histogram of a slice. The *Outputmax* is set at 255 and the *Outputmin* is set to 0. Then

the intensity values of input image pixels are linearly transformed as in equation 4.1. The linear transformation is defined by the minimum and the maximum values that the output image should have and the lower and the upper limits of the intensity window of the input image. Figure 4.1(b) and 4.2(b) show the intensity normalized image. The histogram of the intensity normalized image in Figure 4.1(b) is shown in Figure 4.3(b). The histogram illustrates that the pixel values are normalized between 0–255.

4.3 Initial landmarks detection

As explained in Chapter 2, MRI images are stacked as slices in the axial, coronal and sagittal directions. It is generally observed that, there are always slices that contain no data before the superior, anterior, left or right of the brain appears. The number of empty slices varies with the MRI settings. The superior or top, anterior or front, left and right of the MR volume approximately identify the boundaries of the brain. We detect them as our initial landmarks. The idea is to extract the middle slice in the axial, coronal and sagittal directions by detecting corners in them. The middle slice in each direction approximately represent the dimensions in the whole volume.

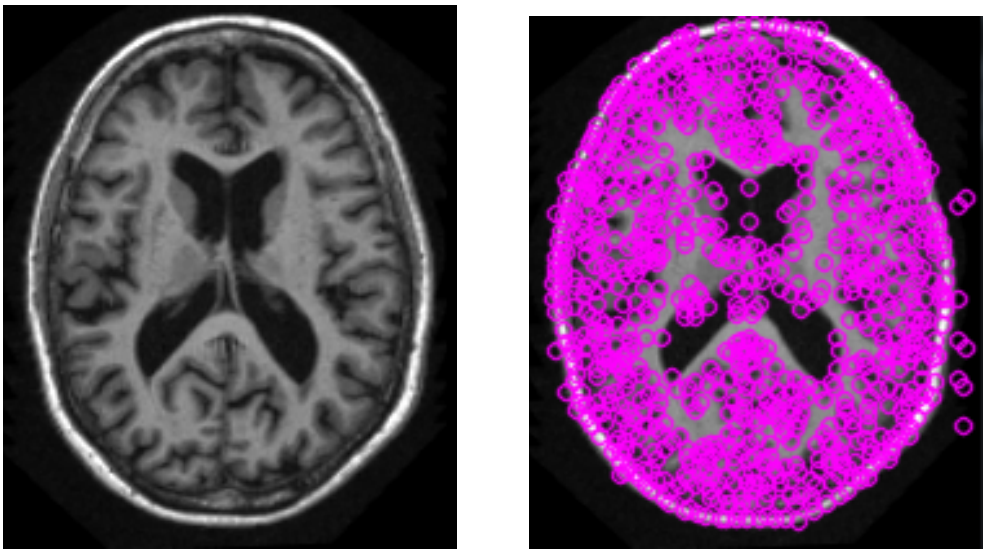


(a) (b)

Figure 4.4: FAST corner detection in the sagittal direction.

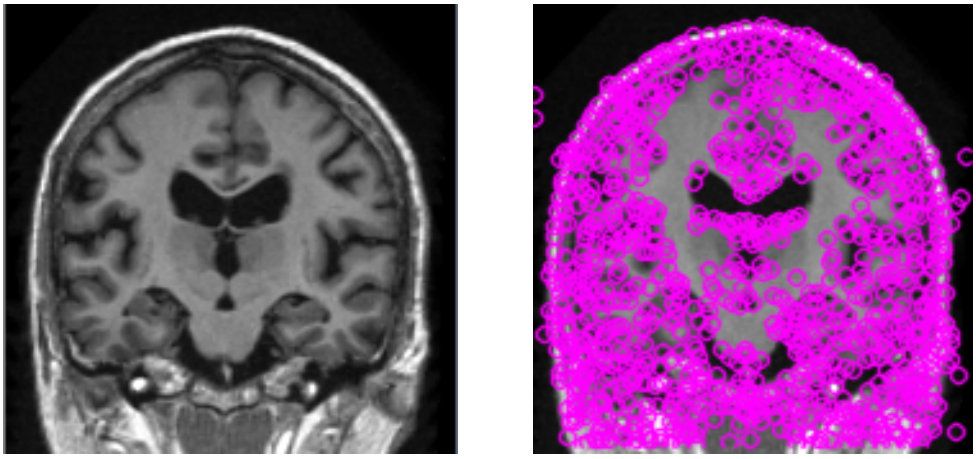
FAST corner detection technique is applied to the middle slices to detect corners. Figure 4.4, 4.5 and 4.6 illustrate the corners detected by FAST in the sagittal, axial and coronal slices. The figures show that a number of corners is detected in the skull of human brain in all three directions. The detected corners are examined line by line until 'n' corners are visited from the origin of an image. The top left hand corner is defined as the origin for an image.

To identify the superior of the brain, the middle slice in the sagittal direction is extracted. For instance the sagittal slice in Figure 4.4(a) is taken and FAST corner detection technique is applied. Figure 4.4(b) depicts the corners detected by FAST in the slice. The corners are examined line by line until 'n' corners are visited. The n^{th} corner is visited approximately at a distance of 2mm inside the skull. To this distance a margin x mm is added to identify the superior of the brain. Similarly to locate the anterior of the brain, the middle slice in the axial direction is considered to which FAST corner detection technique is applied. Figure



(a) (b)

Figure 4.5: FAST corner detection in the axial direction.

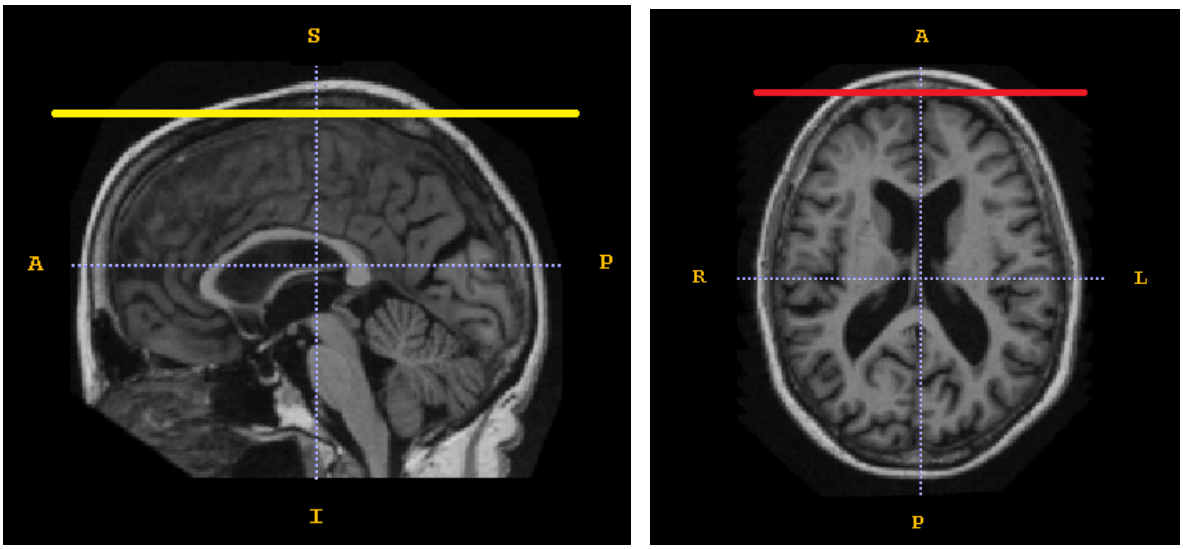


(a) (b)

Figure 4.6: FAST corner detection in the coronal direction.

4.5(a) and 4.5(b) shows the axial slice and the corners detected by FAST. The corners are examined line by line until 'n' corners are visited and a margin x mm is added to locate the anterior of the brain. Figure 4.7(a) shows the yellow line indicating the identified superior landmark and the Figure 4.7(b) shows a red line indicating the anterior landmark of the MR volume.

To identify the left and the right of the brain, the middle slice in the axial direction shown in Figure 4.5(b) is considered again. The corners are analyzed starting from $(0, \frac{h}{2})$ to locate the right and $(w, \frac{h}{2})$ to locate the left of the brain where w and h are the width and height of the slice, instead of the origin. Similar to the anterior and superior landmarks, the corners



(a) (b)

Figure 4.7: Initial landmarks. The markings A, P, R, L, S and I refer to anterior, posterior, right, left, superior and inferior of the brain MRI volume. (a) Superior landmark. (b) Anterior landmark.

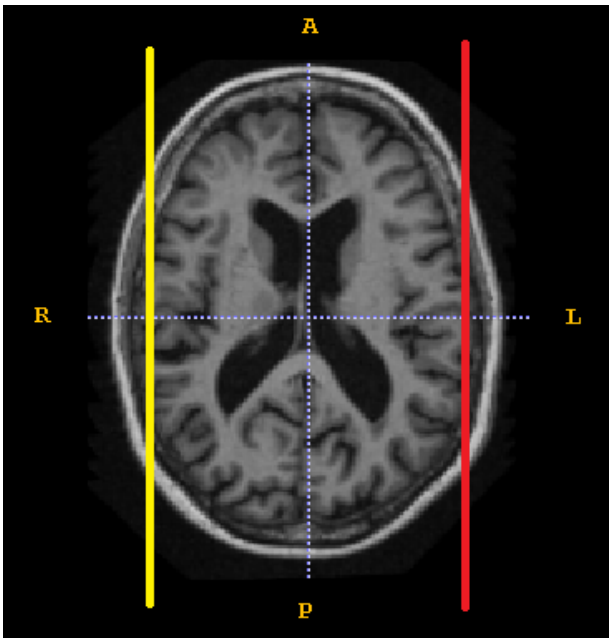


Figure 4.8: Initial landmarks. Left and right of human brain in a MRI volume

are examined line by line till 'n' corners are visited. Figure 4.8 shows the vertical yellow and red line, indicating the left and right of human brain in the MR volume.

The obtained MRI boundaries approximately identify the initial volume of interest that contains the brain, excluding the background.

Chapter 5

SURF Feature Extraction, Description, and Matching

In this chapter a brief explanation on SURF feature extraction, description and matching are discussed. Section 5.1 explains the details of SURF feature extraction. In Section 5.2, SURF feature description is discussed. Later in Section 5.3, we discuss how SURF features are matched.

SURF was proposed by Bay et al. [43]. SURF approximates or even outperforms previously proposed schemes with respect to repeatability, distinctiveness, and robustness, yet is computed and compared much faster. This is achieved by relying on integral images[45] for image convolutions; by building on the strengths of the leading existing detectors and descriptors and by simplifying these methods to the essential. The major stages in SURF are:

- Feature extraction
- Feature description
- Feature matching (when comparing two images.)

5.1 Feature extraction

SURF feature extraction or feature point extraction is based on Hessian matrix because of its good performance on computation time and accuracy. SURF uses integral images [45] to reduce the computation time. SURF relies on the determinant of Hessian for selecting the location and scale [43]. Given a point $x = (x, y)$ in an image I, the Hessian matrix $H(x, \sigma)$ at position x and at a scale σ is defined as follows

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{y_y}(x,\sigma) \end{bmatrix} \tag{5.1}$$

where $L_{xx}(x, \sigma)$ is the convolution of the second order derivative $\frac{\partial^2 x}{\partial x^2} g(\sigma)$ with the image I at position x, and similarly for $L_{xy}(x, \sigma)$ and $L_{y_y}(x, \sigma)$. Gaussians are optimal for scale-space analysis, however the Gaussians needs to discretised and cropped, and even with Guassian

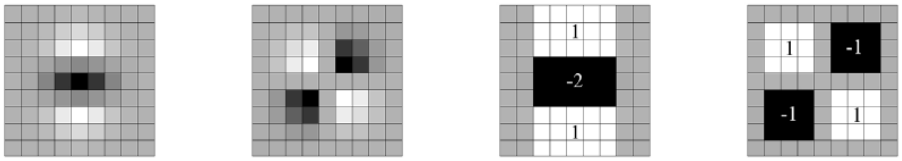
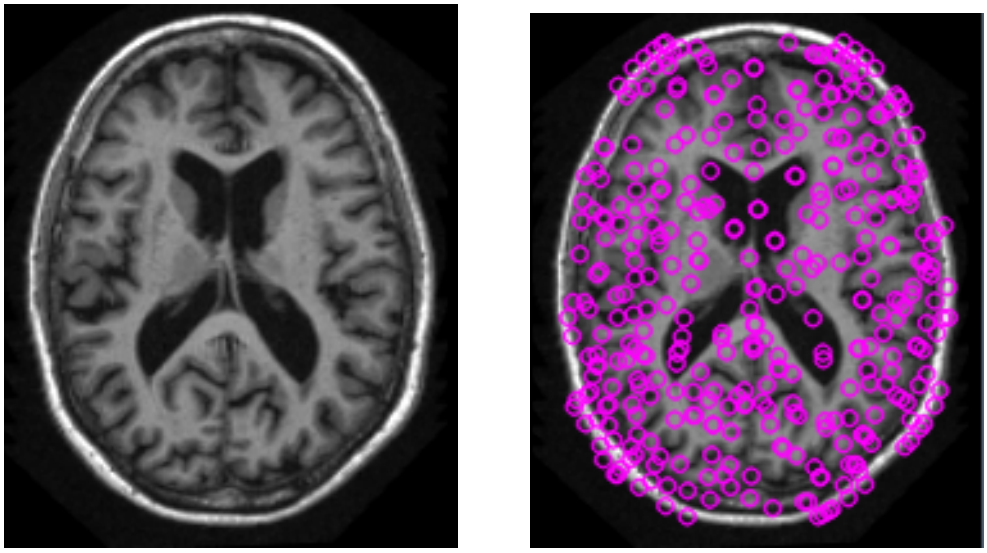


Figure 5.1: Left to right: Gaussian second order partial derivatives in y-direction and xy-direction, and SURF approximations thereof by using box filters. The grey regions are equal to zero.[43]



(a) (b)

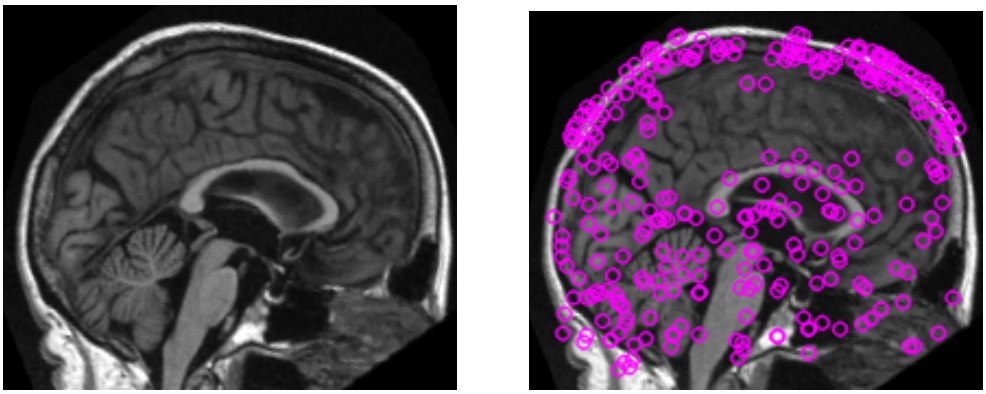
Figure 5.2: SURF feature point detection in the axial direction.

filters aliasing still occurs as soon as the resulting images are sub-sampled. Also, the property that no new structures can appear while going to lower resolutions may have been proven in the 1D case, but is known to not apply in the relevant 2D case [46]. Hence, the importance of the Gaussian seems to have been somewhat overrated in this regard, and SURF tests for a simpler alternative. Similar to LoG approximations [44], SURF uses box filter [43]. These approximate second order Gaussian derivatives, can be evaluated very fast by using integral images, independently of size.

The 9×9 box filters in Figure 5.1 are approximations for Gaussian second order derivatives with $\sigma = 1.2$ and represent lowest scale or highest spatial resolution for SURF. The approximations are denoted by D_{xx}, D_{xy} and D_{y_y} [43]. The weights applied to the rectangular regions are kept simple for computational efficiency. The relative weights are further balanced in the expression for Hessian's determinant with $\frac{|L_{xy}(1.2)|_F|D_{xx}(9)|_F}{|L_{xx}(1.2)|_F|D_{xy}(9)|_F} \simeq 0.9$, where $|x|_F$ is the Frobenius norm. This yields

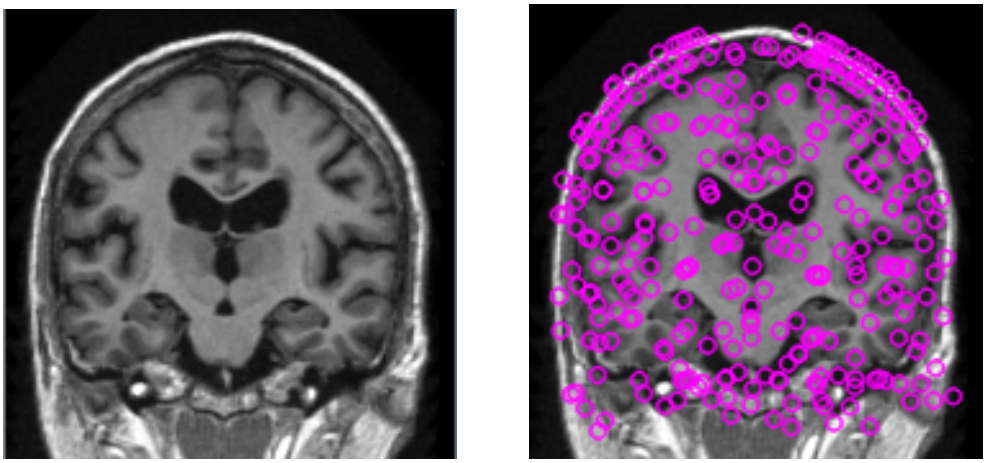
$$det(H_{approx}) = D_{xx}D_{y_y} - (0.9D_{xy})^2 \tag{5.2}$$

Scale spaces are usually implemented as image pyramids. The images are repeatedly smoothed with a Gaussian and subsequently sub-sampled in order to achieve a higher level of



(a) (b)

Figure 5.3: SURF feature point detection in the sagittal direction.



(a) (b)

Figure 5.4: SURF feature point detection in the coronal direction.

the pyramid. Due to the use of box filters and integral images, it is not required to iteratively apply the same filter to the output of a previously filtered layer, but instead can apply such filters of any size at exactly the same speed directly on the original image. Therefore, the scale space is analysed by up-scaling the filter size rather than iteratively reducing the image size [43]. The output of the above 9 × 9 filter is considered as the initial scale layer, which is referred as scale s = 1.2. The following layers are obtained by filtering the image with gradually bigger masks, taking into account the discrete nature of integral images and the specific structure of our filters. Specifically, this results in filters of size 9 × 9, 15 × 15, 27 × 27, etc. At larger scales, the step between consecutive filter sizes is also scaled accordingly. Hence for each new octave, the filter size increase is doubled [43].

As the ratios of filter layout remain constant after scaling, the approximated Gaussian derivatives scale accordingly. Thus, for example, the 27 × 27 filter corresponds to $\sigma = 3 \times 1.2 = 3.6 = s$. In order to localise feature points in the image and over scales, a non-maximum

suppression in a 3 × 3 × 3 neighbourhood is applied. The maxima of the determinant of the Hessian matrix are then interpolated in scale and image space. Figure 5.2, 5.3 and 5.4 show the feature points detected on a slice in the axial, sagittal and coronal directions of a MR dataset respectively. A representative feature set showing the selected scales is shown in figure 5.6.

5.2 Feature description

SURF descriptor is found in two steps [43]:

a) The first step consists of fixing a reproducible orientation based on information from a circular region around the feature point.

b) Then, a square region aligned to the selected orientation is constructed, and the SURF descriptor is extracted from it.

These two steps are explained below.

5.2.1 Orientation Assignment

To be invariant to rotation, a reproducible orientation is identified for the feature points. For this purpose, Haar-wavelet responses are calculated in x and y direction shown in Figure 5.5, and this in a circular neighborhood of radius 6s around the feature point, with s the scale at which the feature point is detected [43]. The wavelet response is calculated at that scale s. Accordingly, at high scales the size of the wavelets is big. Therefore, integral images are used for fast filtering. Only six operations are needed to compute the response in x or y direction at any scale [43].

Once the wavelet responses are calculated and weighted with a Gaussian ($\sigma = 2.5$s) centered at the feature point, the responses are represented as vectors in a space with the horizontal response strength along the abscissa and the vertical response strength along the ordinate. The dominant orientation is estimated by calculating the sum of all responses within a sliding orientation window covering an angle of $\frac{\Pi}{3}$. The horizontal and vertical responses within the window are summed. The two summed responses then yield a new vector. The longest such vector lends its orientation to the feature point [43].

5.2.2 Descriptor Components

For the extraction of the descriptor, a square region centered around the feature point is constructed, and oriented along the orientation described above. The size of this window is 20s [43]. The region is split up regularly into smaller 4 × 4 square sub-regions. This keeps important spatial information in. For each sub-region, few simple features are computed at 5×5 regularly spaced sample points. Haar wavelet responses are calculated in the horizontal d_x and the vertical direction d_y defined in relation to the selected feature point orientation. To increase the robustness towards geometric deformations and localisation errors, the responses are first weighted with a Gaussian ($\sigma = 3.3$s) centered at the feature point. Then, the wavelet responses d_x and d_y are summed up over each subregion and form a first set of entries to the

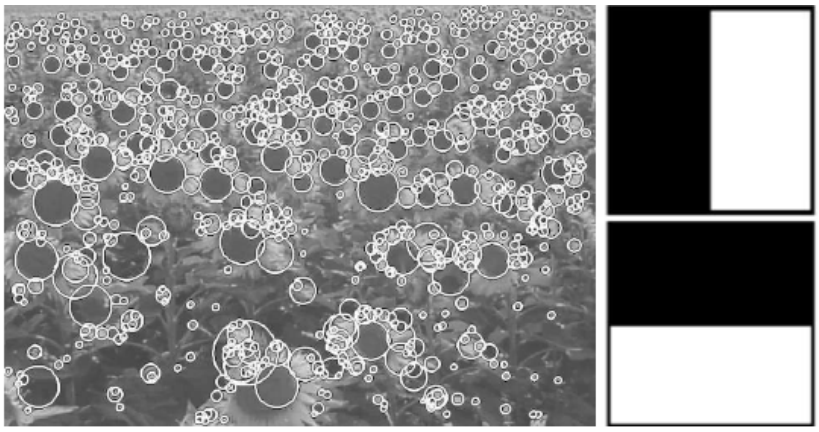


Figure 5.5: Left: Detected feature points for a sunflower field. Right: Haar wavelet types used for SURF [43].

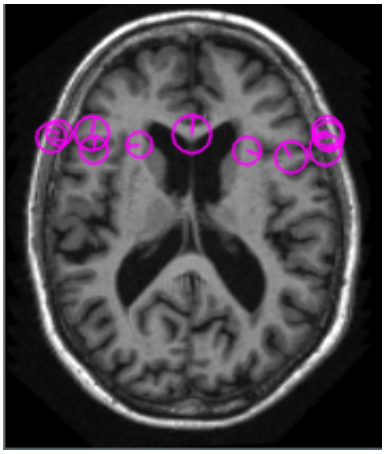


Figure 5.6: Representative sample of scale and orientation assignment to SURF feature points. The circles indicate the scale at which the feature points were detected. The line within the circle indicate the orientation of the feature point.

feature vector. In order to bring in information about the polarity of the intensity changes, the sum of the absolute values of the responses, $|d_x|$ and d_y are extracted which for the second set of entries to the feature vector. Hence, each subregion has a four-dimensional descriptor vector for its underlying intensity structure. This results in a descriptor vector for all 4×4 sub-regions of length 64 [43]. The wavelet responses are invariant to a bias in illumination (offset). Invariance to contrast (a scale factor) is achieved by turning the descriptor into a unit vector.

5.3 Feature matching

The features extracted from two or more images can be tracked in similar images by matching their feature description vectors. Hence the features and their descriptions can be used in a variety of applications such as automated panorama stitching [47], image retrieval [48], object

localisation [44] and in image registration.

Given two images and their features, the matching can be defined by the following equation

$$\forall_{k,l} \sum_{i=1}^{64} d(I_1(P_k^i), I_2(P_l^i)) \tag{5.3}$$

$$d(u, v) = \left(\sum_i (u_i - v_i)^2 \right)^{\frac{1}{2}} \tag{5.4}$$

where I_1 and I_2 are the reference and test images, P_k and P_l are the k^{th} and i^{th} feature points and i is the index of the feature description vector. d is Euclidean distance given by equation 5.4 where u and v are the feature description vectors. In order to match the landmarks between two MRI volumes, for instance a slice in the axial direction, the matching is done as follows. A slice from the test MRI volume is compared to the reference slice from the template MRI volume, by matching their feature points. A feature point in the test image is compared to a feature point in the reference image by calculating the Euclidean distance given by the equation 5.4, between their feature description vectors.

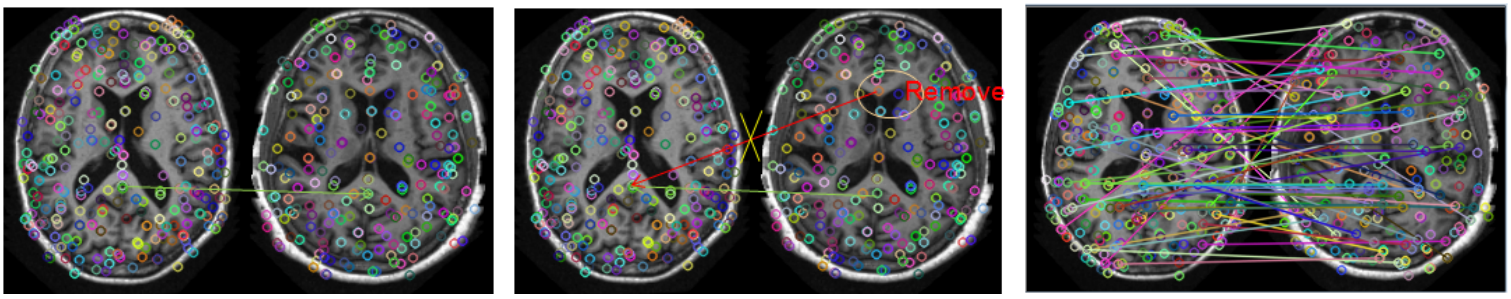
Each feature description vector from the slice in test MRI volume is then compared to all feature description vectors of the reference slice from the template MRI volume. The pair that obtains the best score (that is, the lowest distance between the two vectors) is then kept as the best match for that feature point. This process is repeated for all feature points obtained from the test MRI volume.

For the matches to be robust and reliable, matching is performed both ways, i.e., test slice to reference slice and vice versa. Also two matches are chosen for every feature point based on the best score for every feature point that is observed in the other. Therefore, for each feature point, there are two candidate matches from the other. These candidate matches represent the best matches for a feature point but they do not always correspond to the same landmark. This because of the similarities between brain tissues and non brain tissues. Hence to remove the outliers from the initial candidate matches, the following steps are performed.

5.3.1 Ratio test

The chosen candidate matches are the two best matches based on the best score for every feature point that is observed in the other. If the measured distance is very low for the first candidate match, and much higher for the second candidate match then the first match can be safely accepted as the good one, since it is unambiguously the best choice. Reciprocally, if the two best candidate matches are relatively close in distance, then there exists a possibility that an error is made while selecting one or the other. In this case, both matches are rejected. This is similar to the Nearest-Neighbor Distance Ratio (NN-DR) thresholding technique proposed by Lowe [39]. The accept criteria is given by

$$\frac{d1}{d2} > Threshold \tag{5.5}$$



(a) Matches 1→2 (b) Matches 2→1 (c) Symmetric matches

Figure 5.7: Symmetry test (a) A feature point from reference slice is matched to a test slice. (b) Two feature points from test slice matches to the same feature point in reference slice. The feature point that is not in agreement with both the slice is removed. (c) Symmetric matches after removing outliers.

where $d1$ and $d2$ are the distances to the nearest and 2nd nearest match. The matches that do not satisfy equation 5.5 are removed from the initial list of matches. This is done for each slice in both direction i.e test slice to reference slice and vice versa.

5.3.2 Symmetry test

The ratio test explained in Section 5.3.1 eliminates a number of ambiguous matches. Now there are two relatively good match sets one from the test slice to reference slice and the other from reference slice to the test slice. From these sets, the matches that are in agreement with both sets are extracted. This is the symmetrical matching scheme imposing that, for a match pair to be accepted, both points must be the best matching feature point of the other [49]. Figure 5.7 shows an illustrative example of the matching scheme. Figure 5.7(a) shows a feature point from the reference slice is matched to a feature point in the test data. Similarly, Figure 5.7(b) shows two feature points from the test data are matched to the same feature point in the reference slice. It is observed that one of the two matches in test data is in agreement with both the slices. Hence this matched feature is selected as the best matching feature point of the other, while the second match from test data is discarded. Figure 5.7(c) shows the final matches after the symmetry test.

5.3.3 RANSAC test

The symmetry test explained in Section 5.3.2 removes a good number of outliers. But from Figure 5.7(c) it can be observed that a number of outliers still exists after ratio and symmetry tests. In order to determine the good matches from the list of available matches, an epipolar constraint is used. If the corresponding point of a point p (expressed in homogenous coordinates) is p', and if F is the fundamental matrix between the two slices, then since p' lies on the epipolar line Fp, given by:

$$p'Fp = 0 \tag{5.6}$$

The fundamental matrix associated with an image pair can be estimated from the feature point matches. For the fundamental matrix to be exact, the match set must be made of only

good matches. However, in a real context, it is not possible to guarantee that a match set obtained by comparing the descriptors of detected feature points will be perfectly exact. This is why a fundamental matrix estimation method based on the RANSAC (RANdom SAmpling Consensus) strategy is employed. The RANSAC algorithm aims at estimating a given mathematical entity from a data set that may contain a number of outliers. The idea is to randomly select some data points from the set and perform the estimation only with those points. The number of selected points is the minimum number of points required to estimate the mathematical entity. In the case of the fundamental matrix, eight matched pairs is this minimum number. In this project, OpenCV implementation is used to find fundamental matrix that takes RANSAC method as an input parameter.

Once the fundamental matrix is estimated from the random 8 matches, all of the other matches in the match set are tested against the epipolar constraint that derives from this matrix. All of the matches that fulfill this constraint (that is, matches for which the corresponding feature is at a short distance from its epipolar line) are identified. These matches form the support set of the computed fundamental matrix. Figure 5.8 shows the matches after RANSAC test. It is observed that 95% of the outliers are removed.

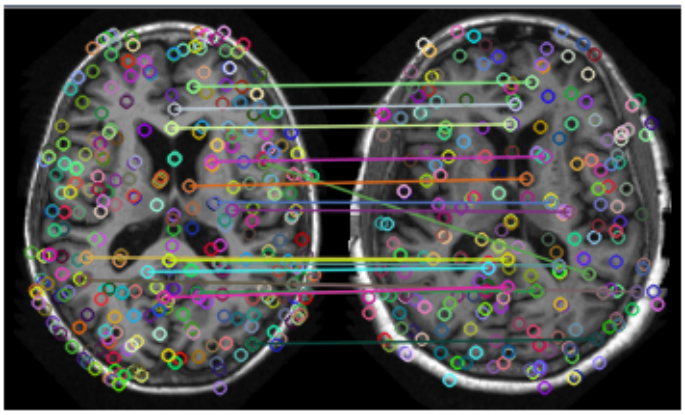


Figure 5.8: Result of RANSAC after removing outliers.

5.3.4 Spatial filtering

This is the final step in obtaining the best inliers from the list of available matches after the RANSAC test. This is a novel contribution to the work presented. It can be noted from the Figure 5.8 that the remaining outlier can be filtered by means of their spatial position(coordinates). For instance a feature point in reference represented by the coordinate $i = (x, y)$, is matched to a feature point with coordinate $i' = (x', y')$ in the test data. A difference of the spatial position of $i - i'$ gives a good indication of their correspondence. The match is accepted if the difference is less than the specified threshold. For slices having the same scale, this can be directly applied to remove the outliers. When the scale of the slices vary, then the spatial position of the features is also scaled in accordance to a scale factor. The scale factor is the ratio of scales of test slice to the reference slice. The threshold is also varied to accommodate the scale changes. The matches are filtered according to the select criteria in equation 5.7. Figure 5.9 shows the result of matched features after spatial filtering.

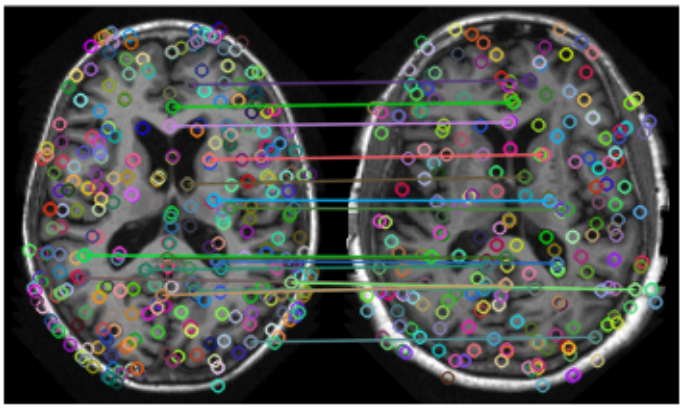


Figure 5.9: Spatial Filtering. The matches are filtered by their coordinates removing the outliers.

$$SelectCriteria = \left\{ \begin{array}{ll} (i - i') < \text{Th} & \text{if scale is equal} \\ (i - i') < \text{s} \times \text{Th} & \text{if scale is different. } s \text{ is the ratio of scales} \end{array} \right\} \quad (5.7)$$

Chapter 6

Analysis of brain MR Images to identify reference landmarks

6.1 Proposed analysis scheme

The feature extraction and matching described in Chapter 5 is used to analyze the brain MR images. The MR image of brain is a volumetric data that contains information in axial, coronal and sagittal directions containing point (voxel), slices and volume. Each of these can be a reference landmark of interest. But not all reference landmarks can be reliably identified in the test MR volumes. Hence we aim to find the reference landmarks that can be reliably identified in most MR volumes.

Let R_1, R_2,..., R_n be the input reference landmarks. The input reference landmarks can either be a structure (voxel, slices and volume) of interest specific to a disease or well known landmarks defined in human brain anatomy, annotated by an expert(s). The input reference landmarks are represented and described by means of their local features points. For instance a point landmark can be represented either by one feature or by features representing a rectangle or circle around the point landmark. A slice landmark is represented by a number of features either in 2D or in 3D(in case of group of slices/volume). The input reference landmarks defined in terms of features are then matched in test MR volumes T_1, T_2,..., T_k. For each input reference landmark R there are k feature match values. Hence for n input reference landmarks, there are $n \times k$ feature match values. The prominence of each reference or the feature match values are then computed by calculating the salient score. The landmarks with high salient score over all tests set are selected. The result of such an analysis is a collection of identified reference landmarks that can be reliably identified in most test MR volumes. In this way the framework learns the landmarks that can be reliably identified.

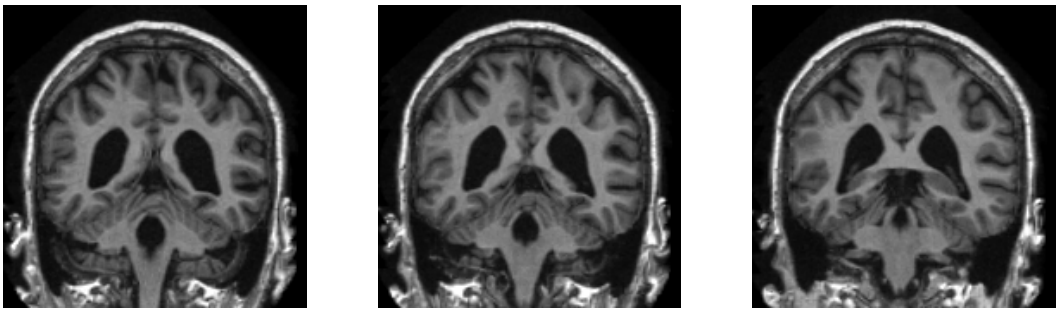
In this project, we are interested in slice landmarks as explained in Chapter 2. The following sections will brief about the experimental setup and the statistical analysis involved to identify the reference slice landmarks.

6.2 Identification of reference slices

To identify the reference landmark slices, 11 test MR volumes having same scale are used. The test MR volumes are registered data set i.e., the correspondence between slices among the test MR volumes are approximately the same, taken from OASIS [50] public archive. For instance an axial slice extracted at the same position from two test MR volumes approximately represent the same structure in the brain. From the test MR volumes, one MR volume is chosen as the reference and the remaining 10 MR volumes are matched against the reference volume. For each input reference landmark a feature match profile is constructed. The profile is a plot of Slice Number Vs. Number of matched feature points. Hence for every input reference landmark there are 10 feature match profiles. This is done for each input landmark taken in axial, coronal and sagittal directions.

Since the structures in the brain are volumetric, it is important to choose landmarks that best represent the brain without redundancy. For instance choosing a landmark slice in the coronal direction at a distance of 2mm from the current selection may represent the same structure in the brain as shown in Figure 6.1. A landmark slice chosen at distance of 5mm from the current selection, show discriminative structures as depicted in Figure 6.1(c). Landmark slice chosen at distance of 10mm shows high variation in the structures resulting in loss of valuable information, so it was not considered. Hence a sampling rate of 5mm is introduced to remove this redundancy. So if the size of MR volume is 176×208×176mm, the input reference landmark is taken at 5, 10, 15... mm.

The process is repeated till each of the tests set is made as reference at-least once. The information extracted from this process is raw and has to be analyzed statistically to derive meaning-full information. The following subsections explain the statistical analysis involved in selecting the reliable slices.



(a) Initial (b) 2mm (c) 5mm

Figure 6.1: Sampling rate. Coronal slice taken at 2mm show similar structures whereas slice taken at 5mm show different structure.

6.2.1 Slice salient score computation

The extracted information contains feature match value i.e. number of matched features for each sample taken in the axial, coronal and sagittal directions. To understand the prominence of the match value, a salient score is calculated for each profile. The salient score is defined

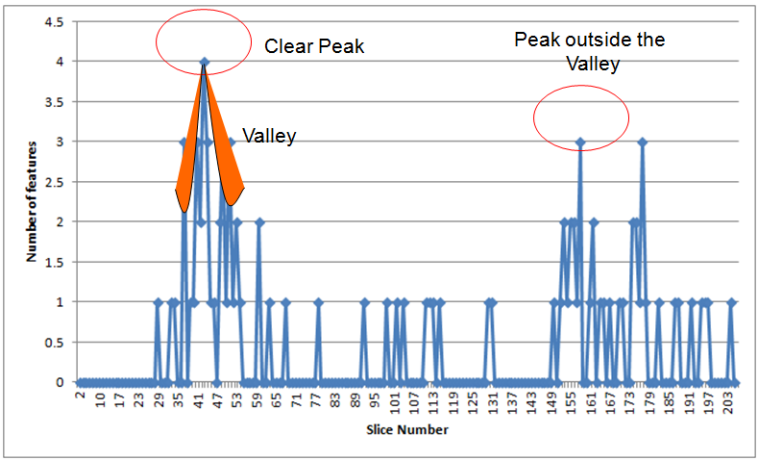


Figure 6.2: Salient score. The valley is marked in orange. The clear peak and the peak outside the valley is marked in red. The ratio of clear peak to the peak outside the valley gives the salient score for a slice.

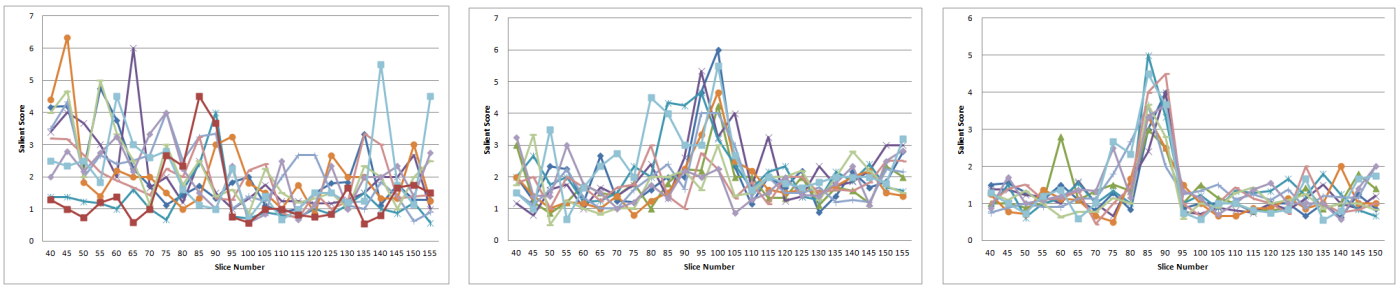
as the ratio of the clear peak value to the maximum peak value outside the valley. The valley is defined as a 10mm distance to either side of the clear peak. The peaks within the valley are not considered for calculating the salient score, since they may contain similar structure as the landmark of interest. When the salient score is less than or equal to 1, the match value has low prominence otherwise the match value has high prominence. Hence a profile with lower salient score can be excluded from further analysis.

$$SalientScore = \frac{ClearPeak}{Peakvalueoutsidethevalley} \tag{6.1}$$

Figure 6.2 show the valley and the peaks that are selected for calculating the salient score. The salient score is computed according to the equation 6.1. Figure 6.3 illustrates the plot of Slice number Vs. Salient scores calculated for samples taken in all three direction for a reference MR volume matched against the remaining 10 test MR volume. From Figure 6.3(c) it can be observed that there is a well distinguished peak around 85mm–90mm in the sagittal direction. Hence it is possible to make a decision to select the reliable slices in the sagittal direction. The pattern is similar for other MR volumes in the tests set.

6.2.2 Median salient score computation

From the previous subsection it can be observed that a decision can be made to identify the reliable slices in the sagittal direction, but for the axial and coronal directions the data should be analyzed further. This is done by computing the median of the salient scores for each input reference landmark along the axial and coronal directions. This is repeated for all the test MR volumes. A plot of Slice number Vs. Median salient score is shown in the Figure 6.4. Figure 6.4(a) and 6.4(b) show the plot of Slice numbers Vs. Median salient scores in the axial and coronal directions. It can be observed from the figures that there are clear peaks in the axial and coronal directions. Hence it is possible to identify reliable slices in those directions. To validate the identified reliable slices in the sagittal direction in the above subsection, we plot the median salient score for the MR volumes in the sagittal direction. The Figure 6.4(c)



(a) (b) (c)

Figure 6.3: Salient scores for a reference MR volume. Plot of Slice number Vs. Salient score for each input reference slice landmark.(a) Profile in the axial direction. (b) Profile in the coronal direction. (c) Profile in the sagittal direction. The profile in the sagittal direction shows a clear peak around 85mm-90mm.

validates the claim for reliable slices in the sagittal direction showing peaks at 85mm and 90mm.

6.2.3 Reference slice selection

An algorithm is used to select the reliable slices in the axial and coronal directions automatically based on a threshold T defined on median salient score. The algorithm ranks slices that are above the defined threshold T. The flow chart to automatically rank the slices is shown in Figure 6.5. The algorithm follows the following logic:

Step 1 Select the maximum median salient score above T and mark it as the current selection.

Step 2 Assign a priority value for the current selection, the lowest value being the highest priority.

Step 3 Define a valley(5mm) around the current selection and discard the median salient score in the valley for the next iteration.

Step 4 Go to Step 1 and repeat till all the median scores above T are selected at-least once.

The result of such an algorithm over the median salient scores of all the test MR volume is shown in Figure 6.6. The thresholds for axial and coronal directions are set at 2 and 1.5 respectively. The figure gives a good indication of the prominence of slices and their rank of selection by means of their priority value. The peaks are distributed throughout the MR volume in the axial and coronal directions. The peaks can be selected to represent different regions of the brain. For instance in the coronal direction, slices 50, 95 and 125 represent the posterior, middle and anterior region of the brain. Similarly for axial direction slices 55, 85 and 105 represent the inferior, middle and superior region of the brain. Hence these slices can be chosen as reference slices or reference landmarks in the axial and coronal directions, that can be reliably identified in most MR volumes.

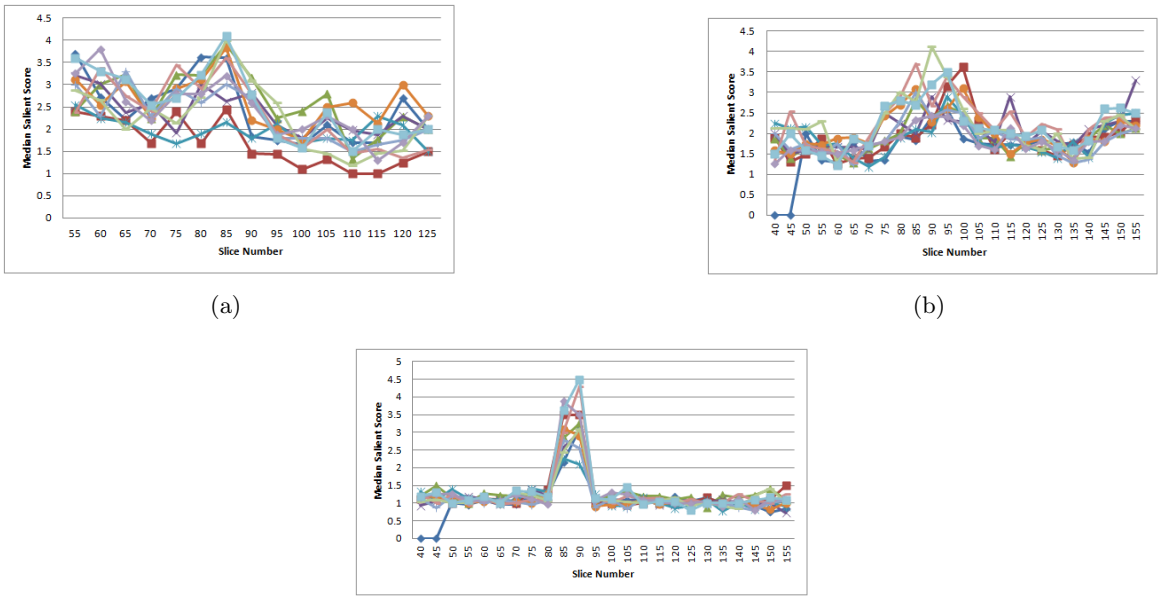


Figure 6.4: Median salient scores for 11 MRI test volume. (a) Plot of Slice number Vs. Median salient score each input reference slice landmark in the axial direction. (b) Plot of Slice number Vs. Median salient score each input reference slice landmark in the coronal direction. (c) Plot of Slice number Vs. Median salient score each input reference slice landmark in the sagittal direction.

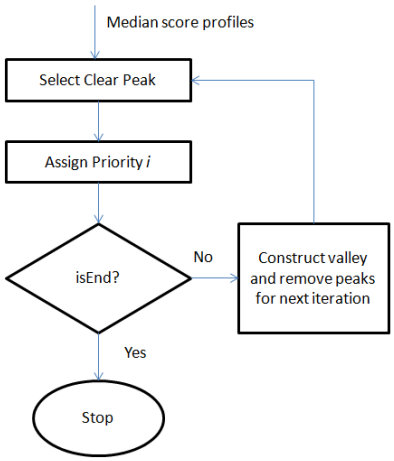
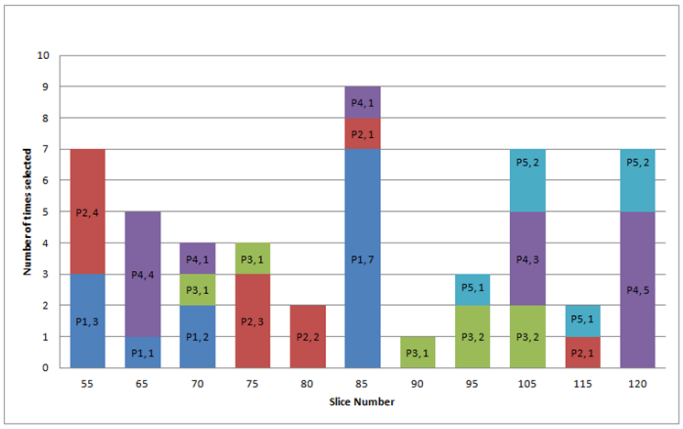
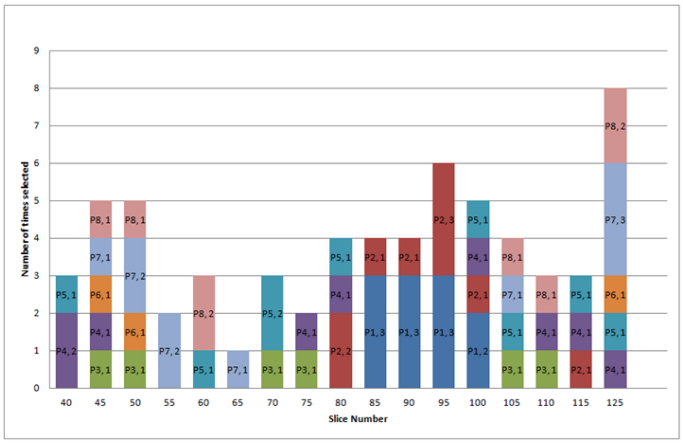


Figure 6.5: Flow chart for automatic ranking of reliable slices.



(a)



(b)

Figure 6.6: Result of automatic ranking of slices. (a) The ranking of slices with their priority in the axial direction. (b) The ranking of slices with their priority in the coronal direction. It is observed that the peaks are distributed throughout the volume. These peaks can be selected as the reference slices.

Chapter 7

Reference based localization of landmarks

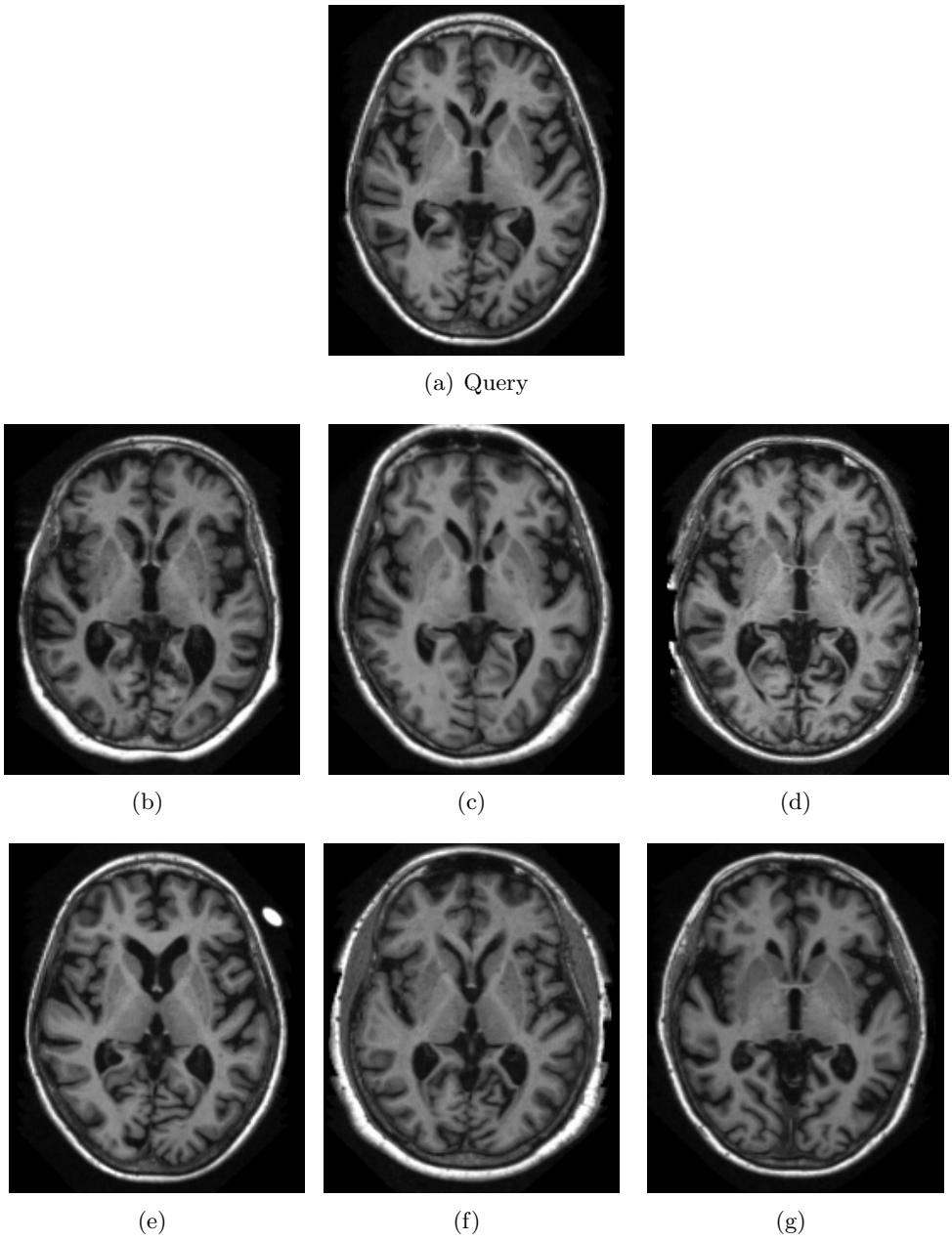
In this chapter, reference based localization of landmarks are discussed. The localization of clinically important landmark in brain images is crucial for many neurological studies such as surgical planning and diagnosis of specific diseases. The localized landmarks can also be used to register an MR volume. Conventional manual landmark annotation requires expertise and is often time-consuming. In this chapter we use the knowledge gained from analysis of MR volume in Chapter 6. We show the gained knowledge can be used to localize landmarks robustly. We also show that the accuracy of the framework increases with the knowledge gained from analysis stage. In Section 7.1, search and retrieval of landmark slices based on the reference landmarks identified in Chapter 6 is discussed. First we present the work without reference landmarks and then with reference landmarks. We show that the accuracy of landmark localization improves with the use reference landmarks. In Section 7.2, matching different MRI volumes or localizing group of slices using reference landmarks is discussed. In Section 7.3, detection of mid sagittal plane (MSP) using reference landmarks is discussed.

7.1 Search and retrieval of landmark

In the medical domain, experts usually look at specific anatomical structures to identify the cause of a pathology, and therefore they can largely benefit from automated tools that retrieve relevant slice(s) from a patients image volume in diagnosis. In diagnosing diseases with high prevalence and unknown cause or progress, medical experts can largely benefit from patient–to–patient search methods that compare multiple patient data and retrieve relevant cases [12]. Accordingly, retrieving the relevant slice given a query, which is a specific case of patient-to-patient search, can be of further help to the expert in diagnosis of anatomical structure specific diseases, such as hypocampus or basal ganglia disorders of the brain [12].

7.1.1 Without reference landmarks

Given a query slice, similar slice(s) are searched and retrieved from a test MR volume. The query slice is represented and described by its features. These features are searched in all slices of a test MR volume by matching them. Each slice in the test MR volume is represented



(a) Query

(b) (c) (d)

(e) (f) (g)

Figure 7.1: Search and Retrieval – same scale. (a) Good input query slice. (b–e) Retrieved slices from 6 MR volumes.

by a match score. For a retrieval task, the test slice with the highest match score is assigned as the most similar slice to the query and retrieved. It is possible to retrieve slices that are similar to the query slice by analyzing the match scores in the test MR volume and ranking them according to the match score.

For example a query slice in the axial direction is shown in Figure 7.1(a). The query slice is around the AC–PC region of the brain has discriminative structures and hence can be

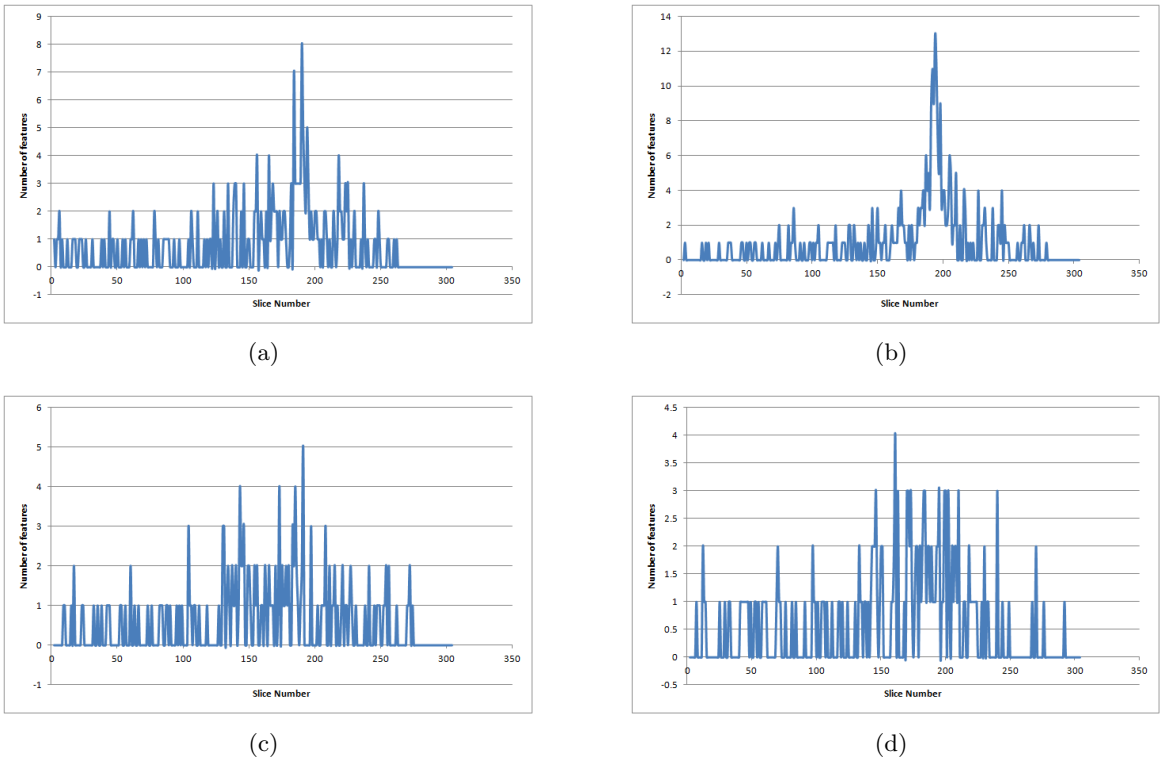
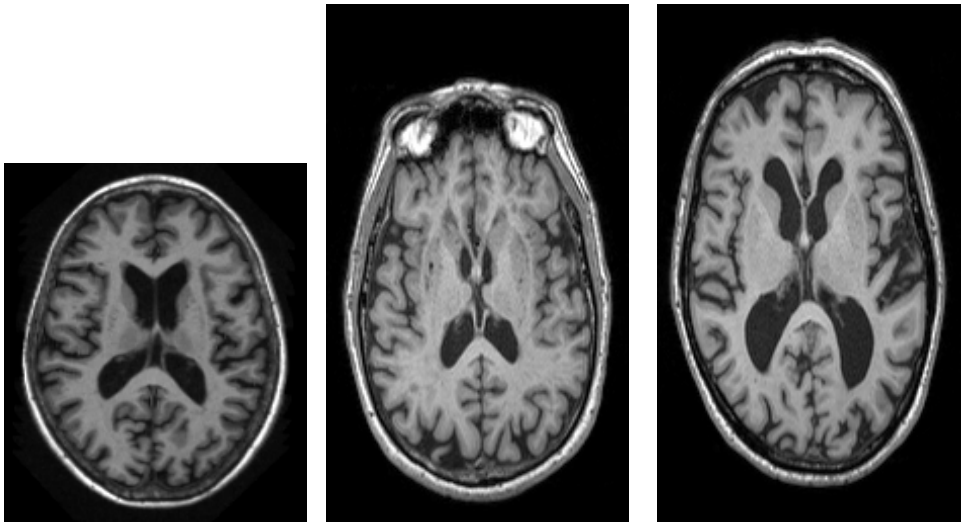


Figure 7.2: Profiles around ventricle region - MR volume with different scales. The profile is a plot of Slice Number Vs Number of matched feature. (a-c) Profiles with clear peak and correspond to same structure as query. (d) Profile showing clear peak which is also false positive.

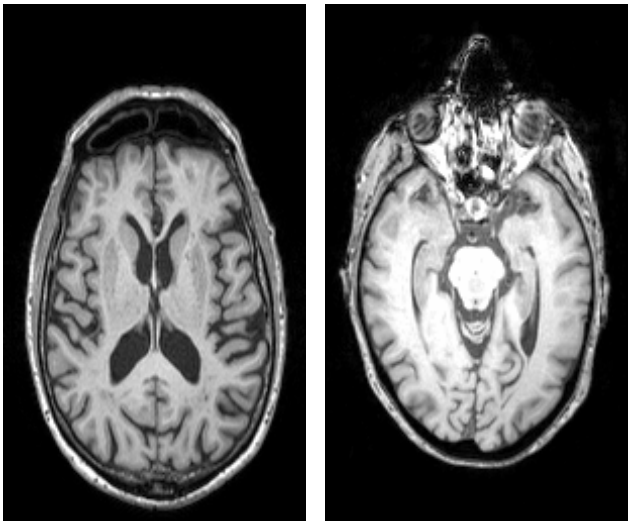
matched easily. Such a slice is considered as a good query. The framework uses this query slice as its input. The framework extracts and matches query features against all slices in the test MR volume. The query slice with the highest number of matches with respect to the test slice is retrieved. Figure 7.1 shows the query and the corresponding retrieved slice from 6 MR volumes having the same scale. Therefore, given a query, the framework automatically retrieves relevant slice(s) from a patient's image volume.

Similarly a slice that has discriminative structures around the ventricular region is manually annotated and selected as a query slice. In this case, the query and the test MR volume have different scales. The query slice is retrieved by matching its features against all slices in the test MR volume. The query slice with the highest number of matches with respect to the test slice is retrieved. Figure 7.2 shows the profile for the four test MR volumes corresponding to the query shown in Figure 7.3(a). The profile is a plot of Slice Number Vs. Match score. Figure 7.3 show their corresponding retrieved slices from the four test MR volume.

However, the query landmarks can vary from patient-to-patient and also depends on diagnosis to specific diseases. Hence, there is a possibility that a given query may not have discriminative structures and hence not reliably identified by feature extraction and matching



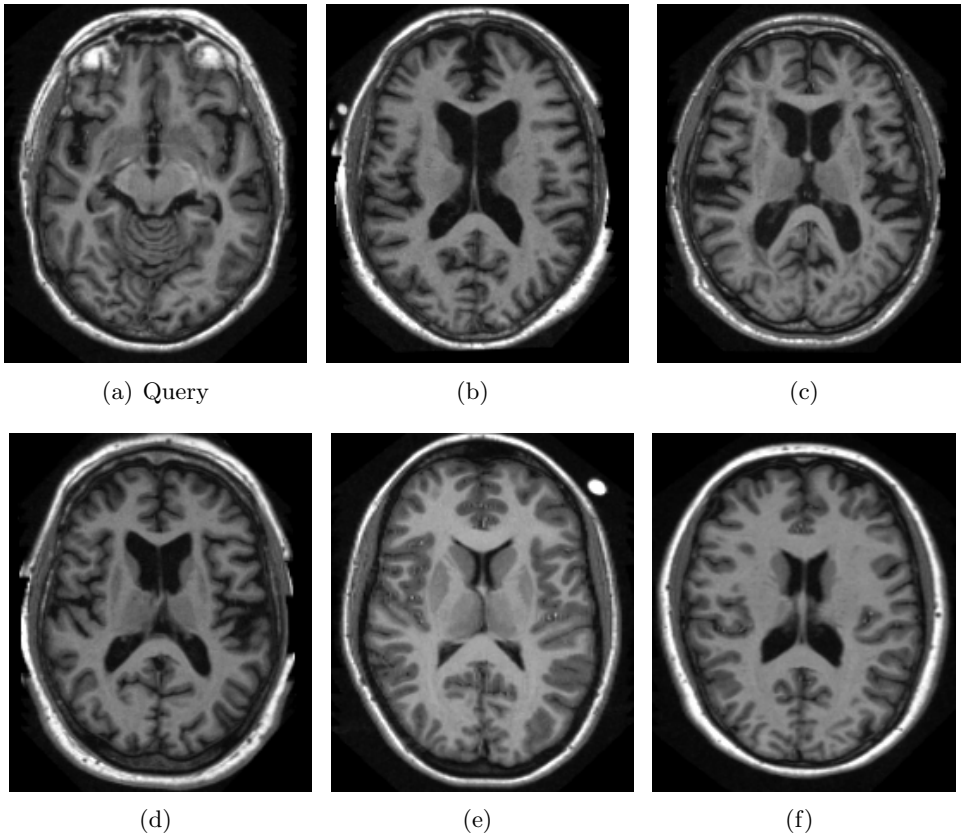
(a) Query (b) (c)



(d) (e)

Figure 7.3: Search and Retrieval – different scales. (a) Good input query slice. (b–e) Retrieved slices from 4 MR volumes.

technique. This might lead to retrieving wrong slices or false positives. A slice that does not have discriminative structure or a bad query slice is shown in Figure 7.4(a). The retrieved slices from five test MR volumes are shown in Figure 7.4(b) – 7.4(f). The figures show that the retrieved slices are not the intended target slice in the test MR volume but are false positives. False positive is defined as the ratio of number of wrongly retrieved slices to the total number of query. Table 7.1 shows the percentage false positive among 10 test MR volume having the same scale for the query slice in Figure 7.4. The table also show the mean error for correctly retrieved slices(True Positive(TP)), wrongly retrieved slices(False positive (FP)) and the mean error over 10 test MR volumes. For true positives the mean error is small but



(a) Query (b) (c)

(d) (e) (f)

Figure 7.4: Search and Retrieval - Without reference slices. (a)Bad query slice. (b-f) Retrieved slices from five MR volumes showing false positives.

is considerably high for false positives. Sometimes wrong slice or false positives are observed even when the query slice have discriminative structures. The retrieved slice in Figure 7.3(e) is a false positive. Hence given a query slice, it is possible that the retrieved slice is a false positive. The current methods to search and retrieve landmarks from an MR volume employ similar techniques and hence they are prone to error when the query is not discriminative enough.

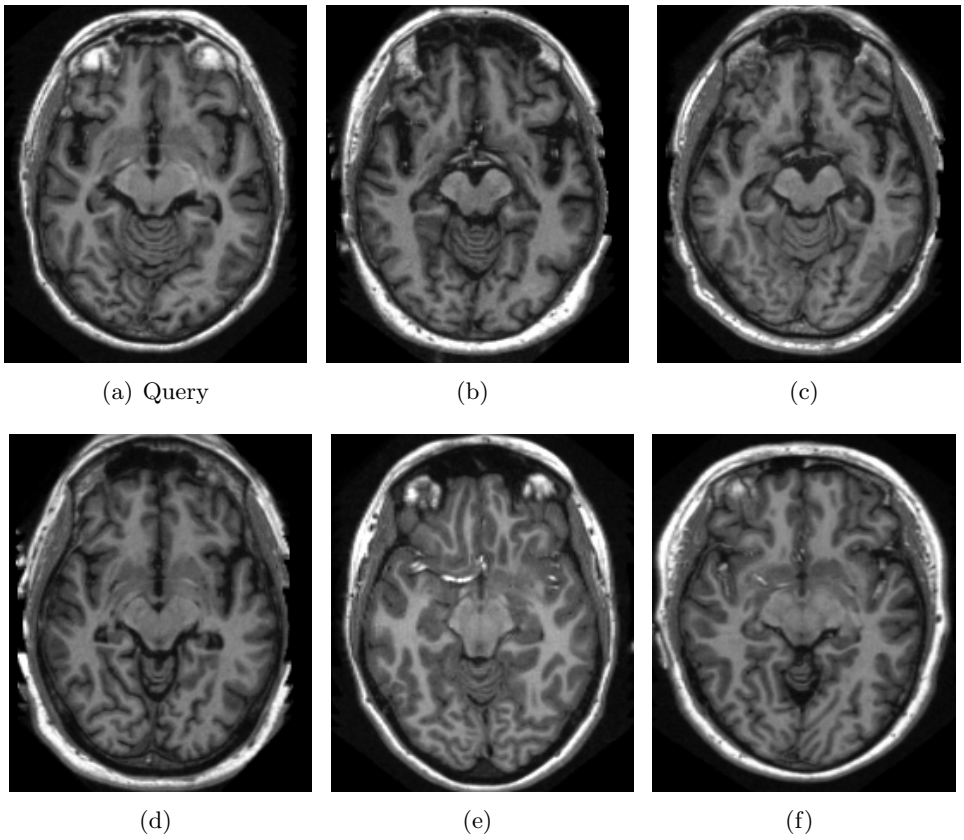
7.1.2 With reference landmarks

The search and retrieval technique explained in the earlier subsection can retrieve wrong slices when the given query can not be reliably identified in a test MR volume. To overcome this problem and to identify any query slice, we propose to use the reference slices identified in Chapter 6 and the initial landmarks identified in Chapter 4. The framework initially localizes the reference slices in the test MR volume through feature extraction and matching. The localized slices in the test MR volume together with the initial landmarks demarcate brain into different regions. The localized slices in the test MR volume also give an approximate correspondence between the reference volume and the test volume. The query slice is then

Table 7.1: Mean error and percentage of false positive for the query slice in Figure 7.4 among 10 MR volumes.

% FP	TP Mean Err. (mm)	FP Mean Err. (mm)	Mean error (mm)
40	1.16	23.09	12.12

localized to one of the regions in the test MR volume say R, using the position of the query slice in the reference MR volume. The position of the query slice is then scaled within the region to locate the query slice in the test MR volume. The scaling can be linear or non-linear based on the scales of the reference and test MR volume. This is the initial estimate P of the query slice that is to be retrieved from the test MR volume.

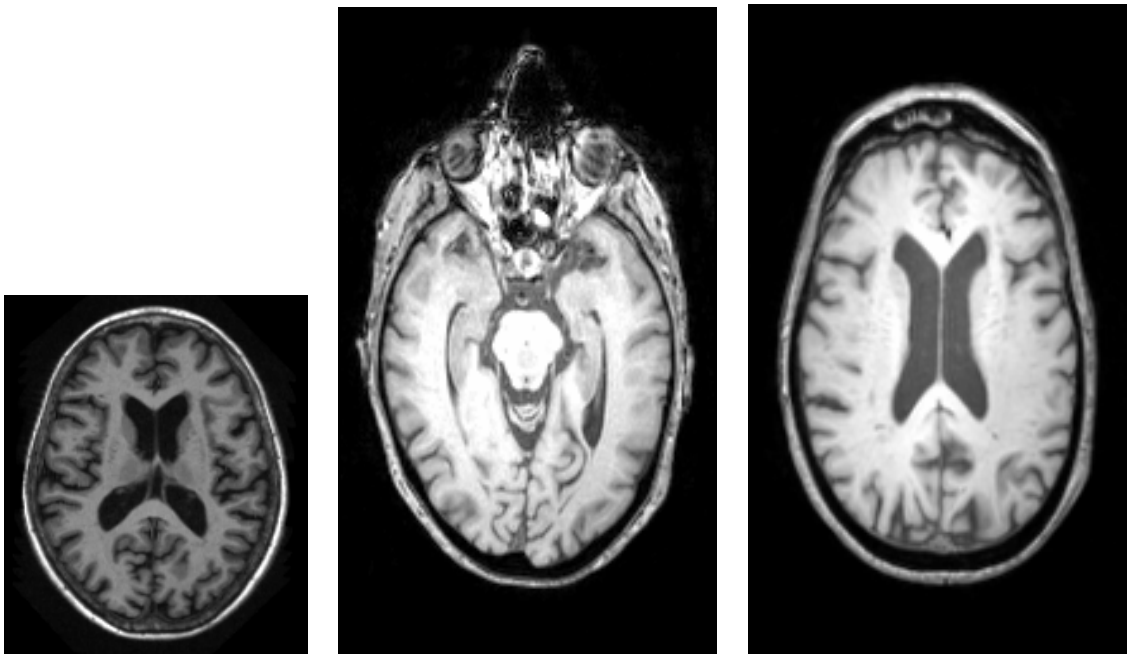


(a) Query (b) (c)

(d) (e) (f)

Figure 7.5: Search and Retrieval - With reference slices. (a) Bad query slice. (b-f) Retrieved slices from five MR volumes showing correct slices.

To increase the accuracy and robustness of the search and retrieval method, the query slice is searched independently of the reference slices as mentioned in the previous subsection. The profile is analyzed and the position of the clear peak is selected as the second estimate P' of the query slice in the test MR volume. The final position of the retrieved slice is found according to the equation 7.1.



(a) Query (b) (c)

Figure 7.6: Search and Retrieval - With reference slices.(a) Good input query (b)Retrieved slice without reference landmarks is 40mm from the ground truth. (c)Retrieved slice with reference landmarks is 3mm from the ground truth.

$$Retrieveposition = \left\{ \begin{array}{ll} P & \text{if } P' \text{ lies outside R} \\ P' & \text{if } abs(P - P') < 5mm \\ avg(P, P') & \text{if } 5mm \leq abs(P - P') < 10mm \\ P & \text{if } abs(P - P') \geq 10mm \end{array} \right\} \quad (7.1)$$

where P and P' are position estimates with reference slices and without reference slices.

When the position estimate P' falls outside the localized region R, then P is selected. If the position estimates fall within the localized region R and they are less than 5mm apart, then the position estimate P' is selected. If the position estimates are less than 10mm apart, then the average of their positions is selected. When the estimated positions are more than 10mm apart, then P is selected. The slice at this position is the final retrieved slice.

Using the reference landmarks improves the accuracy and reliability of the search and retrieval operation. Given the same query slice in Figure 7.4(a), the false positives are reduced considerably with the use of reference slices. Figure 7.5 shows the retrieved slices for the query slice observed in the previous subsection. The retrieved slices show similar structures as the query slice. Table 7.2 show the search and retrieval results with reference slices with the same query as in Figure 7.4. The search is done on the same 10 test MR volumes chosen in the previous subsection. The results show good improvement in the mean error of the retrieved slice for true positives(TP), false positives(FP) and for the whole search and retrieval

application. The percentage of false positives are also considerably reduced with the reference slices.

Table 7.2: Mean error and percentage of false positive for the query slice in Figure 7.4 among 10 test MR volumes - with reference slices

% FP	TP Mean Err. (mm)	FP Mean Err. (mm)	Mean error (mm)
10	0.3	3	1.65

In cases where the input query slice is discriminative enough, but shows false positive as in Figure 7.3(e), the reference landmarks removes the false positives and improves the detection accuracy as shown in Figure 7.6.

7.2 Volume matching

Volume matching identifies group of slices rather than a single slice. Such an application can be used to study the correspondence between two MR volumes or the output of the application can be an input to applications such as detection of point landmarks and segmentation of brain structures. The requested group of slices is given as an input to the framework by its spatial position in the reference MR volume. The framework identifies the requested volume in two steps. First, by localizing the reference slices identified in Chapter 6 in the test MR volume which gives an approximate correspondence with the reference MR volume and identify the approximate group or volume of interest by scaling the positions of the requested group of slices. Second, the approximate volume of interest is fine tuned by matching the requested group of slices with the approximate group of slices. The profile with the highest match score is selected as the final volume of interest. The two steps are explained below.

First, the framework localizes the reference slices in the test MR volume through feature extraction and matching. The localized slices identify the correspondence between the test and reference MR volumes. Then the approximate group of slices is retrieved either by linear or non-linear scaling, based on the scales of reference and test MR volumes. The scaling is linear when both the template and test MR volumes are of same scale and non-linear otherwise. The requested group of slices is represented by their spatial positions in the reference MR volume and hence they can be localized in a region either between the reference slices or between the MR boundaries and the reference slices. So, the requested group of slices can be localized within a region in the test MR volumes using the localized references and the MR boundaries identified in Chapter 4. The lower and the upper bound of the requested group of slices is found by its position in the reference MR volume. These bounds are used to identify the approximate volume of interest with in a localized region in the test MR volume. For volumes with the same scale, linear scaling is performed by adding the difference between the reliable slice in the reference and the volume bound i.e, the lower and the upper bounds, to the corresponding identified slices in the test MR volume.

The scaling factor for non-linear scaling is defined as the ratio of the distance between the identified slices in the test MR volume to the distance between the reliable slices in the template MR volume. The scaling factor varies according to the distance between the

reference slice identified in the test MR volume and vary non-linearly. For scaling factor computation, the initial landmark identified in Chapter 4 is used. For instance the axial direction is divided into five regions. The region is a group of slices between the identified slices in the test MR volume. Suppose the position of reference slices in the axial direction are R_1, R_2 and R_3, and the position of localized slices in the test MR volume are S_1,S_2 and S_3 and the initial landmarks CS and CI represent the position of superior and inferior in the test MR volume. Then the axial direction can be demarcated into four regions having four scaling factor, one for each region given by the following equations:

$$S1 = \frac{S_1 - S_2}{R_1 - R_2} \tag{7.2}$$

$$S2 = \frac{S_2 - S_3}{R_2 - R_3} \tag{7.3}$$

$$S3 = \frac{S_3 - CS}{R_3 - Templatesuperiorposition} \tag{7.4}$$

$$S4 = \frac{S_1 - CI}{R_3 - Templateinferiorposition} \tag{7.5}$$

where the current superior(CS) and inferior(CI) positions are identified as initial landmarks in Chapter 4 and they vary for each test MR volume. Thus in the axial direction, the bounds are scaled non-linearly according to equations 7.2–7.5. Similarly for the coronal and sagittal directions the scaling factor is determined by localizing the reference slices in the test MR volume.

Second, the approximate group of slices are fine tuned to identify the final group of slices. A profile for the requested group of slices is constructed by matching the boundary and middle slices of the requested region to the approximate group of slices. This is repeated within a window of $5mm$. The lower bound and the upper bound of the approximate group of slices is modified by $\pm 5mm$. The profile with the highest match value is chosen and retrieved as the final group of slices.

7.3 Mid sagittal plane detection

In this section, we describe another application of the proposed framework. Identification of mid sagittal plane is important since they are used in a variety of application such as AC–PC point landmark detection, registration and to study brain symmetry. The proposed framework can be easily used to identify the MSP in two ways. First, by providing any slice as an query to the proposed frame. Second, by using the identified reference landmarks in the sagittal direction from analysis stage. These methods are explained below.

In the sagittal direction, feature extraction and matching can be used to identify the symmetry of brain hemispheres, that is a slice with similar anatomical structure can be identified in the brain hemispheres. Given an input query slice, feature extraction and matching is performed against all slices in a test MR volume in the sagittal direction. The slices with the highest match score is selected. Figure 7.7 shows a profile of a slice landmark in the sagittal

direction depicting the symmetry observed in brain hemispheres. It can be observed that there are two peaks approximately at the same distance from boundary of each hemisphere. The Figure7.8 shows the query and the retrieved slices from a test MR volume. Let the position of the retrieved slices be P and P'. The average of P and P' gives the approximate position of MSP. To reliably identify the MSP position, average position estimates from a number of sagittal query slices can be taken.

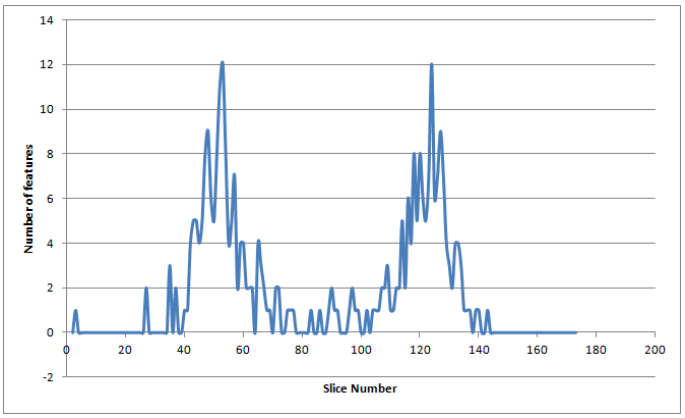
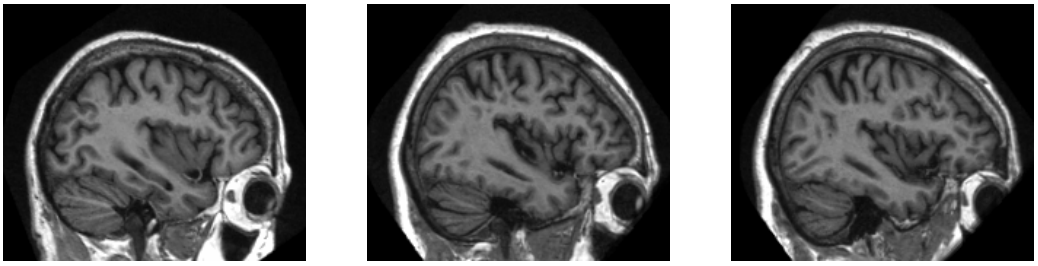


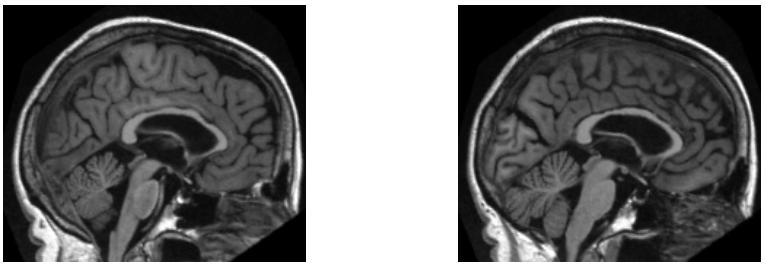
Figure 7.7: Profile of a slice in the sagittal direction. The profile shows peaks at two region around slice number 50 and 125, depicting the symmetry of brain hemispheres in the sagittal direction.

Secondly, we use the reference landmark slices identified for the sagittal direction in the analysis stage. The identified reference slices slices are shown in the Figure 7.9. The reference slices are matched in a test MR volume. The positions of the clear peak are identified. The profile of a test MR volume matched against the reference is shown in Figure 7.10. The clear peaks are observed at 84mm and 87mm. The distance between peaks of the profiles is the region containing the MSP. In the profiles shown in Figure 7.10 the distance between the peaks is $3mm$. This $3mm$ region can be an input to other applications to detect the exact MSP location. Alternatively, the average of the positions in the $3mm$ region can be chosen as the position of mid sagittal plane. This is more reliable and robust approach compared to the first approach, since the reference slices are reliably identified in most MR volumes.



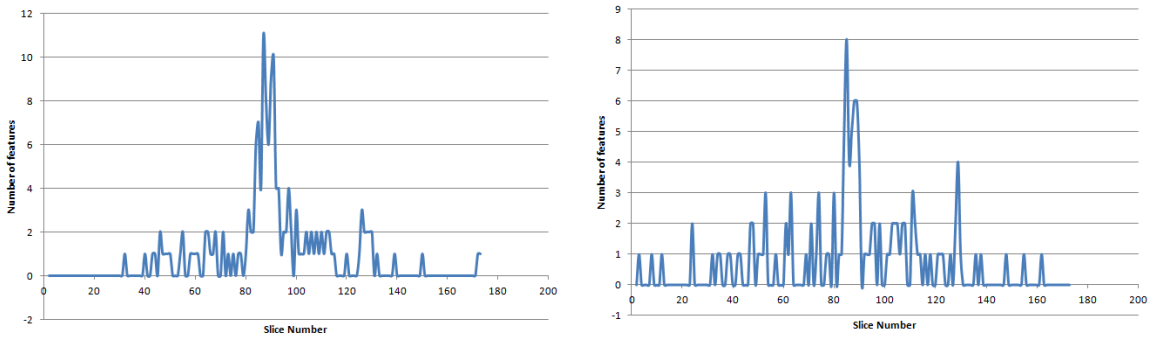
(a) Query (b) (c)

Figure 7.8: Brain symmetry. (a) Query slice. (b-c) Retrieved slices from a MR volume.



(a) (b)

Figure 7.9: Reference slices in sagittal direction. (a) Reference slice 1. (b)Reference slice 2.



(a) (b)

Figure 7.10: MSP detection using reference slices. (a-b) Profiles showing a clear peak and low peak values outside the valley for an MR volume. The difference between the peak values identify the region within which the MSP is contained.

Chapter 8

Results

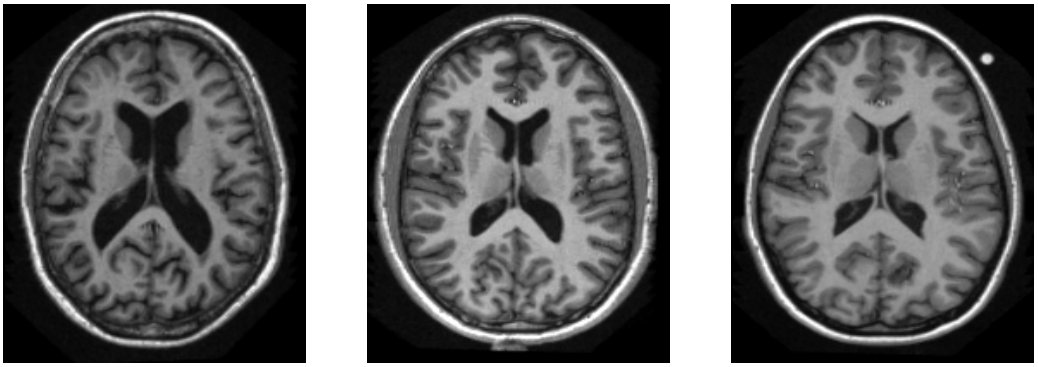
In this chapter the quantitative and qualitative results of this project are discussed. In Section 8.1 the details about the test dataset and their respective sources are explained. In Section 8.2, the test setup for the experiments is discussed. In Section 8.3, the ground truth generation for the tests set is explained. In Section 8.4, the results about slice matching and the performance of framework are discussed. In Section 8.5, the results of volume matching and its accuracy is presented.In Section 8.6, the performance of the framework with SURF and SIFT is discussed.

8.1 Image data

The database comprises of 470 MR T1 and 14 T2 weighted images. Out of which 435 MRI volumes of size $176 \times 208 \times 176$mm are taken from OASIS public archive [50], provided under an open access data use agreement. The subjects are all right-handed and include both men and women. 100 of the included subjects over the age of 60 have been clinically diagnosed with very mild to moderate Alzheimers disease. The remaining 35 MRI volumes of size $150 \times 256 \times 256$mm are taken from IXI public archive [51]. The 14 T2 weighted images of size $256 \times 256 \times 130$mm are taken from IXI public archive [51].

8.2 Test setup

The algorithm was implemented in C++ using OpenCV [49] and ITK [52] on a Windows 7 OS with intel i5 processor. OpenCV is used to extract and match features while ITK was used for preprocessing the MRI volume. Three MRI volumes are selected to represent the different age groups and they act as reference. The reference MR volumes are taken from the OASIS dataset for T1 weighted images and from IXI for T2 weighted images. Out of the 470 T1 weighted images, 60 are selected alphabetically to form the T1 test set. 30 of which are chosen from the OASIS dataset and the remaining 30 are taken from IXI dataset. The tests set are chosen such that the included MR volumes did not contribute to the analysis done in Chapter 6. Similarly out of the 14 T2 weighted dataset, 3 MR volumes are selected as reference and the remaining 11 are made the test set.

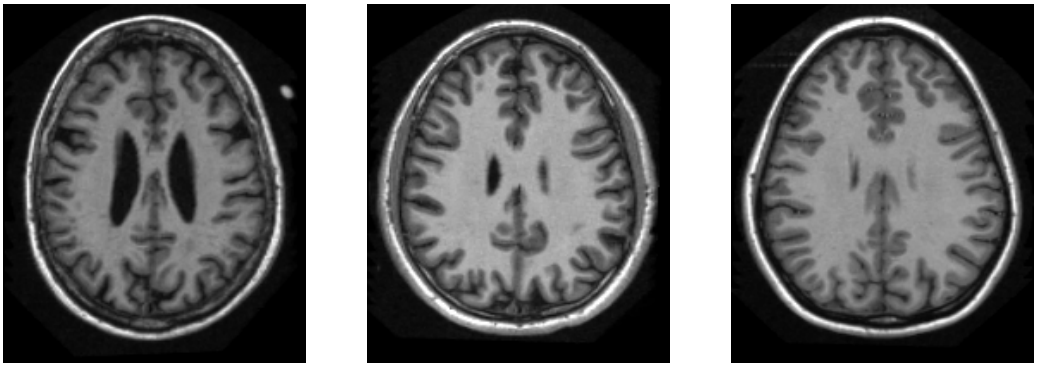


(a) (b) (c)

Figure 8.1: Selected reliable slice in the axial direction in 3 references - Reference slice 1

8.3 Ground truth generation

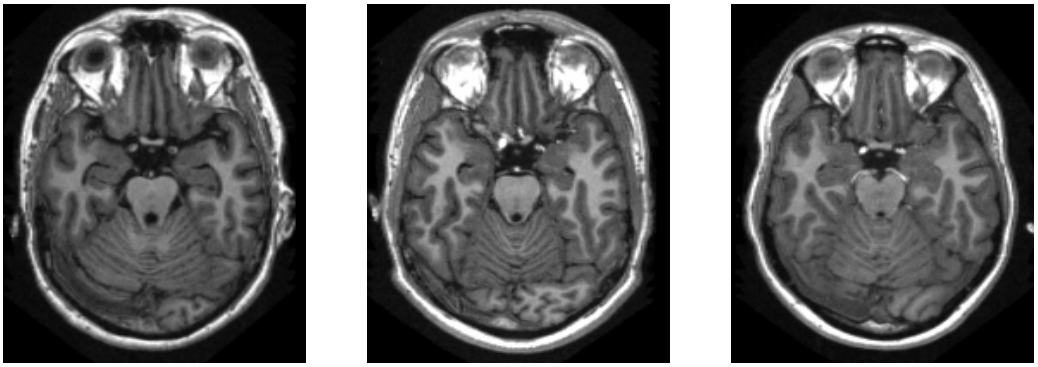
The ground truth for all 71 cases, 30 OASIS and 30 IXI T1 weighted dataset and 11 T2 IXI dataset, were provided independently by two researchers by manually annotating the reference landmark slices identified in Chapter 6, in the test dataset. Figure 8.1, 8.2, 8.3 ,8.4 and 8.5 show the selected reference slices in the axial and coronal directions. The Euclidean distance in millimeters between the ground truth and the position obtained by the algorithm was the error measure. The mean and standard deviation of the estimation error of landmarks against the ground truth are calculated.



(a) (b) (c)

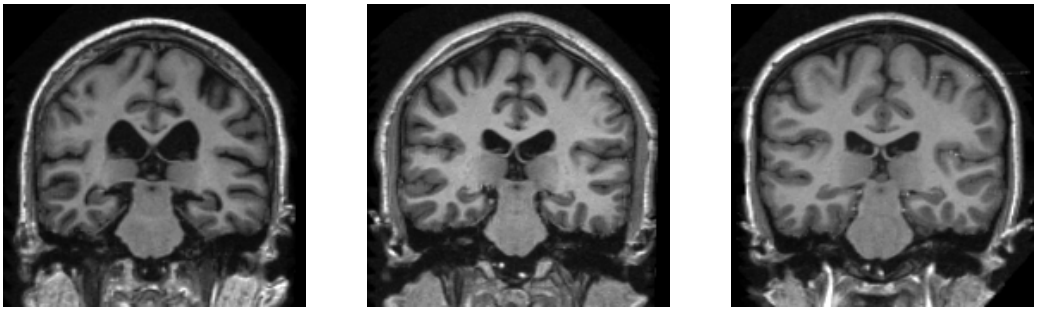
Figure 8.2: Selected reliable slice in the axial direction in 3 references - Reference slice 2

The inter-observer variability analysis was carried out by finding the difference in the positions of reference slices, identified by the researchers. Table 8.1 shows the inter-observer variability as a difference in the positions. In the axial direction, the reference slice in Figure 8.3 has the highest percentage difference. We attribute this to the presence of similar slices around the reference slice. The reference slice in Figure 8.1 has the lowest percentage difference, since the structures are discriminative can be identified by the researchers more easily. In the coronal direction, the reference slices in Figure 8.4 and 8.5 show similar percentage



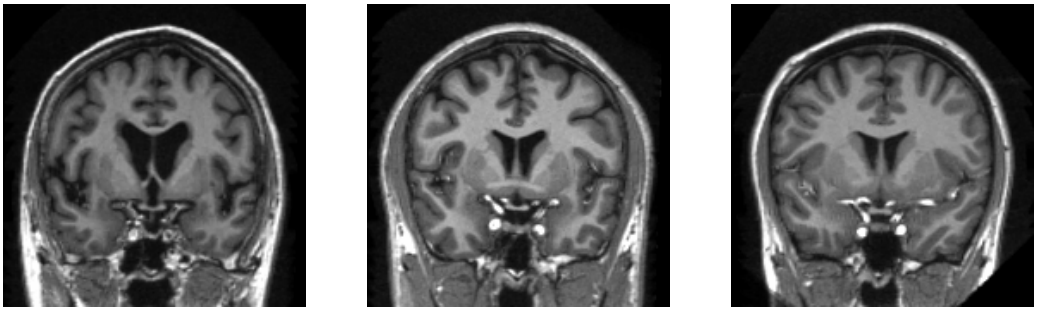
(a) (b) (c)

Figure 8.3: Selected reference slice in the axial direction in 3 references - Reference slice 3



(a) (b) (c)

Figure 8.4: Selected reliable slice in the coronal direction in 3 references - Reference slice 1



(a) (b) (c)

Figure 8.5: Selected reliable slice in the coronal direction in 3 references - Reference slice 2

difference. The inter-observer variability analysis shows that the identified reference slice positions vary based on the researchers perception and when the reference does not show distinct structures compared to its neighbors, the variability between the researchers increases. In this project we took the average of the reference slice positions by the researchers as the ground truth.

Table 8.1: Inter-observer variability between Researcher-1 and Researcher-2 by the difference in the identified position of reference slices.

Direction	Ref. slice 1 diff. (mm)	Ref. slice 2 diff. (mm)	Ref. slice 3 diff (mm)
Axial	1.2	2.8	7
Coronal	1.3	1.2	-

8.4 Results of slice matching

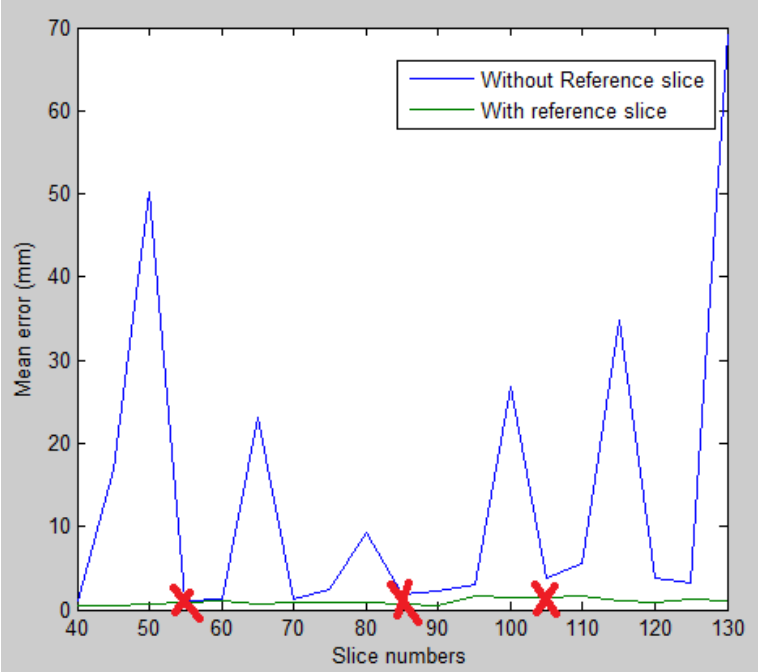


Figure 8.6: Plot of Slice Vs. Mean error in (mm) in 30 OASIS test set in the axial direction. The green and the blue curve show the mean error with and without reference slice. The cross mark indicate the identified reference slices.

First, we present the results of slice matching for T1 weighted dataset. The axial slices from an T1 weighted OASIS MR volume are given as the query slice to the framework. The axial slices are sampled at an interval of 5mm. The axial slices are matched against 30 T1 weighted OASIS test data set chosen in the test setup. Figure 8.6 show the results of slice matching for the corresponding slices. The figure is a plot of axial slice numbers Vs. Mean error in *mm* i.e., the Euclidean distance from the ground truth. The matching is first performed without reference slices and then with reference slice. The results of slice matching without reference slice show a mean error up to 70mm. With reference slices, the query slices are detected accurately and the mean error is approximately the same for all slices. The mean error with reference slices is between 1mm-2mm. Hence, there is a good improvement in the mean error compared to the method without reference slices. The cross mark in the Figure 8.6 indicate the position of the chosen reference slices. It is observed from the figure that at chosen reference slices, the mean error without reference slice converges to the mean error

with reference slice. Table 8.2 shows the mean minimum, mean maximum and median error in mm for 30 T1 weighted MR volumes, without and with reference slices. It is observed that the mean minimum error for both the approaches between 0.5mm and 1.0mm inclusive. But when reference slices are not used, the error increases considerably for query slices that are not reliably identified.

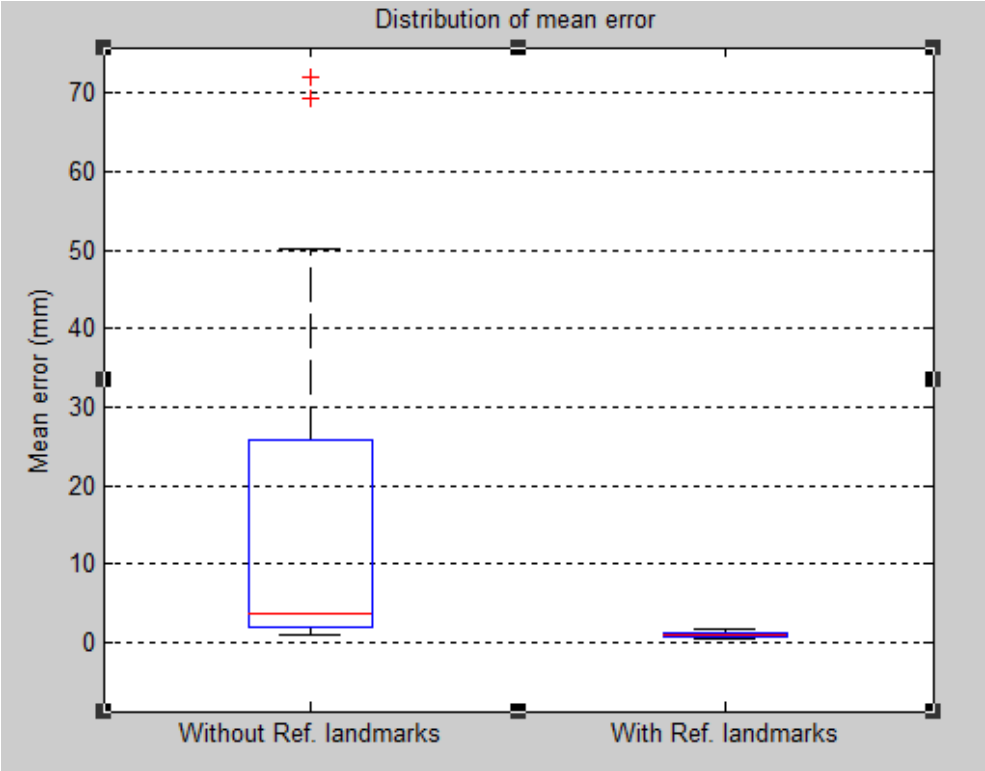


Figure 8.7: Box plot showing the distribution of Mean error in (mm) in 30 OASIS test set in the axial direction.

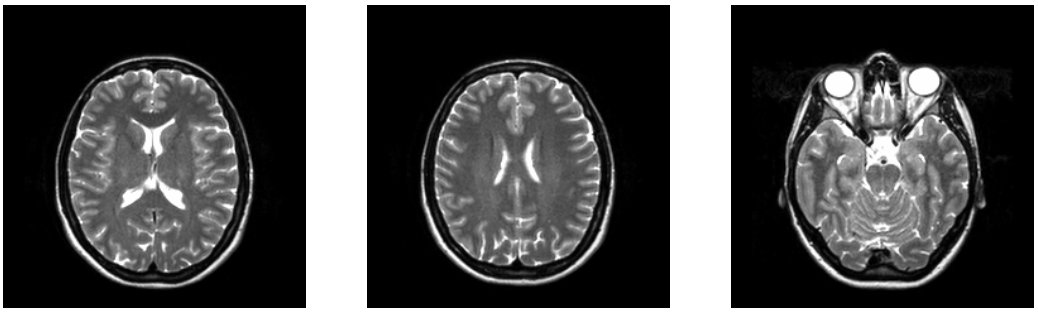
Table 8.2: Distribution of mean error for axial query slices among 30 T1 weighted MR volumes

No: query slice	Mean minimum Err. (mm)	Mean maximum Err. (mm)	Median error (mm)
Without reference slices			
19	1	69.30	3.72
With reference slices			
19	0.50	1.70	0.90

The box plot in Figure 8.7 show the distribution of mean error in mm, when the framework do not use reference slices and when the framework use reference slices. The plots show that the error is high when there are no reference slices and low when there are reference slices.

Similarly for T2 weighted images, the input query landmarks in the axial direction are

shown in Figure 8.8(a), 8.8(b) and 8.8(c). The slice matching results for T2 dataset are shown in Table 8.3. The query slices are matched and identified within a maximum mean error of 1.4mm. Hence, the developed framework is contrast independent and can be applied to both T1 and T2 weighted dataset.



(a) (b) (c)

Figure 8.8: T2 weighted MR images. (a) Query slice 1. (b) Query slice 2. (c) Query slice 3.

Table 8.3: T2 weighted MR volume: Reference slice matching results

Data	No. of datasets	Mean Error (mm)	SD of Error
Volume with same scaling	11	-	-
Query slice 1	11	0.27	0.90
Query slice 2	10	1.34	1.82
Query slice 3	10	1.0	1.77

The accuracy of the framework is measured by the percentage of true positives and false positives. Since the performance of the framework depend on how well the reference slices are identified in the test MR volumes, the true positives and false positives are computed for the same. To find the accuracy at which reference slices are detected, the slices are retrieved from 60 T1 weighted MR volumes. The reference slice is retrieved for all the 60 T1 weighted dataset. The position of the retrieved slice is compared with the ground truth by computing the Euclidean distance in millimeters between them.

Table 8.4 shows the mean error and standard deviation of the retrieved slice in 60 T1 weighted test MR volumes in the axial direction. It is observed that the mean error for all volumes in the axial direction is 0.38mm for reference slice 1, whereas the mean error increases to 1.1mm for reference slice 2. This is because the reference slice 1 has more discriminative structure than reference slice 2, and hence feature extraction and matching identify reference slice 1 more accurately. All reference slices in the axial direction are identified within a mean distance of 1.5mm from the ground truth.

Similarly, Table 8.5 shows the mean error and the standard deviation for reference slices in the coronal direction. The mean error for all volumes in the coronal direction is 0.062mm for reference slice 1 and 0.2 for reference slice 2. The reference slices in the coronal direction

Table 8.4: Mean and SD of error for reference slices in axial direction in 60 T1 weighted MR volumes

Data	No. of datasets	Mean Error (mm)	SD of Error
Reference slice 1			
All volumes	60	0.38	1.57
OASIS	30	0.33	0.88
IXI	30	0.44	2.14
Reference slice 2			
All volumes	60	1.09	2.30
OASIS	30	1.31	2.00
IXI	30	0.84	2.60
Reference slice 3			
All volumes	60	0.82	1.16
OASIS	30	0.70	1.15
IXI	30	0.95	1.20

are identified within a distance 1mm of ground truth in the OASIS tests set whereas for IXI, the reference slices are identified within a distance of 2mm from the ground truth.

Table 8.5: Mean and SD of error for reference slices in the coronal direction in 60 T1 weighted MR volumes.

Data	No. of datasets	Mean Error (mm)	SD of Error
Reference slice 1			
All volumes	60	0.06	2.54
OASIS	30	0.23	0.97
IXI	30	0.62	2.12
Reference slice 2			
All volumes	60	0.20	2.30
OASIS	30	0.14	0.99
IXI	30	0.28	1.31

True positives and false positives are calculated according to the following equation

$$TP = \frac{Numberofcorrectlyretrievedslices}{Totalnumberofinputquery} \times 100 \tag{8.1}$$

$$FP = \frac{Numberofwronglyretrievedslices}{Totalnumberofinputquery} \times 100 \tag{8.2}$$

Correctly retrieved slices are defined as those slices that are identified within an absolute distance of 5mm from the ground truth. All the slices outside this limit are considered false positives. In addition to identifying the true and false positives, the framework makes no decision when there is no sufficient information. This makes it possible for the clinical technician to identify the required structure manually. Table 8.6 show the accuracy of framework in the axial and sagittal directions. The average framework accuracy in the axial direction is 87.7% whereas in the coronal direction it is 78%.

Table 8.6: Framework Accuracy

Slice	No. of datasets	True Positive (%)	False Positive (%)	No decision (%)
Axial				
Reference slice - 1	60	93.33	5	1.67
Reference slice - 2	60	91.6	8.4	-
Reference slice - 3	60	78.4	16.6	5
Coronal				
Reference Slice - 1	60	78	22	-
Reference Slice - 2	60	80	20	-

8.5 Results of volume matching

For volume matching, 15 OASIS dataset and 15 IXI dataset are taken from the 60 T1 weighted dataset. The bounds of the group of slices are given as the input to the framework. The retrieved group of slices is compared with ground truth by their Euclidean distance. The error is calculated by computing the average error of the bounds. Table 8.7 show the mean error for 30 dataset. The framework identifies query group of slices within a 4mm distance in case of OASIS dataset and 6mm in case of IXI dataset.

Table 8.7: Results of volume matching

Data	No. of datasets	Mean Error (mm)
OASIS	15	3.43
IXI	15	5.24

8.6 Performance comparison between SIFT and SURF

To benchmark the results obtained using SURF, we compare the results of the framework with SURF and with SIFT. 10 MR volumes out of the chosen 30 OASIS dataset is taken alphabetically. The reference slices in Figure 8.1, 8.2 and 8.3 were given as an input query slice to the framework. The error is calculated by the Euclidean distance between the retrieved position and ground truth. Table 8.8 shows the mean error of 10 MR volumes by SURF and

SIFT, for reference slices in the axial direction. It is observed that the accuracy of the retrieved slices in both SIFT and SURF are similar. The accuracy of SIFT in detecting reference slice - 3, is slightly higher than SURF but at a higher computation cost. Table 8.9 show the computation time required to retrieve all axial reference slices in an MR volume. It is seen than SURF is approximately 50% faster than SIFT.

Table 8.8: Mean error in the axial direction. SURF and SIFT.

Slice	No. of datasets	Mean error (mm)	SD
SURF			
Reference slice - 1	10	0	0
Reference slice - 2	10	1.6	1.70
Reference slice - 3	10	0.7	1.05
SIFT			
Reference Slice - 1	10	0.2	0.42
Reference Slice - 2	10	1.4	1.70
Reference slice - 3	10	0.1	0.31

Table 8.9: Computation time between SIFT and SURF in the axial direction for an MR volume of size $176 \times 208 \times 176$mm

Method	Time (sec)
SIFT	227.7
SURF	130.7

Chapter 9

Discussion and conclusion

The results of slice matching in Chapter 8 show that the framework detects query landmarks more accurately with reference slices. When reference slices are not used the detection rate is low and have higher error value. The mean error value starts to rise rapidly beyond the ventricle region when reference slices are not used. This is because beyond the ventricular region the slices are not discriminative enough and the background starts to appear. Hence feature extraction and matching do not perform well in this region. With reference slices the landmarks are detected within a mean error of 2mm in both the axial and coronal direction. The accuracy of the framework is highest in the axial direction. We attribute this to the discriminative structures observed in the axial direction. Similarly, for the coronal and sagittal direction we observed that the landmarks are detected more accurately with reference slices. The difference in the mean error value with and without reference slices prove that the accuracy of the framework improves considerably with reference slices.

The framework was tested for contrast independent detection by using T2 weighted MR images. The selected query slices closely resemble the reference slices identified for T1 weighted images. The accuracy of the framework with T2 weighted images is similar or comparable to that of T1 weighted MR images. In some cases the mean error value is lower than that of T1 reference slices. While testing for slice matching, we also included a couple of MR datasets that had orientation difference. The framework was able to correctly retrieve the query slice. The framework was used to detect slice, group of slices or volume and plane landmarks. In principle the framework can be used to detect any landmark in brain. The framework was tested with SIFT and SURF feature extractors. The results in Chapter 8 show that reliable slice 1 in the axial direction was detected more accurately SURF, but reliable slices 2 and 3 in the axial direction were detected accurately by SIFT. We believe by comparing the methods over larger dataset the mean error rate of SIFT and SURF will approximately be similar.

The novel contributions to this thesis work include (a) a generic framework development that can be extended to detect any landmark in brain. (b) Using SURF to extract robust and repeatable features and to the best of our knowledge SURF was never applied on MR brain images. (c) Analyzing the MR volumes using robust features and identifying reference landmarks that can be reliably detected in most MR volumes. (d) Prove that the accuracy of landmark detection improves considerably with reference slices identified by analyzing the

MR volumes. (e) Contrast independent landmark detection i.e. applicable to both T1 and T2 weighted images.

Future works includes (a)point landmark detection. The initial analysis over 100 test MR volumes to identify consistent point landmarks did not provide us with good results. The analysis showed that point landmarks showed a maximum repetition rate of 40% over 100 MR volume. We believe analysing more test MR volumes and using lower sampling rate than 5mm can identify consistent point landmarks and will give better results. (b) The background estimation in this thesis work is performed by calculating the local histogram value for a slice instead of the complete volume and then defining a threshold at 8% of the histogram value to eliminate background. This approach is not adaptive to input MR volume. This can be made adaptive by using the estimates of MR image boundaries. The MR boundaries approximately identify the brain tissues excluding the background. Hence using the intensity distribution of the brain tissues and the local histogram estimate, the background detection can be made adaptive to the input MR volume. (c) The computation time of the framework can be reduced considerably given that feature extraction and matching can be implemented in a GPU. In this project this was not tried since the focus was to use SURF to detect landmarks in brain MR images.

To conclude, in this project, we have developed a framework that can be extended to detect all landmarks in brain. We have used SURF feature extraction and matching technique and applied it successfully on brain MR images. We have analyzed brain MR volumes by extracting robust features and identified reference landmarks by statistical analysis, which can be reliably detected in most MR volumes. We have proved that the accuracy of the framework improves considerably using the reference slices found by analysing the brain MR images. We have shown that the framework can be used to detect slice landmarks, match group of slices or volume landmarks and in mid sagittal plane detection. The performance of the framework with SURF is comparable to that of framework with SIFT in terms of detection accuracy, but at a 50% faster computation time than SIFT.

Bibliography

[1] I. Dryden and K. V. Mardia, *Statistical Shape Analysis.* Wiley, 1998.

[2] W. Nowinski(a), "Computerized brain atlases for surgery of movement disorders," in *Seminars in Neurosurgery*, vol. 12, no. 2, 2001, pp. 183–194.

[3] W. Nowinski(b), "Electronic brain atlases: features and applications. in: 3d image processing: Techniques and clinical applications (eds. caramella d, bartolozzi c)," in *Medical Radiology series, Springer-Verlag*, 2002, pp. 79–93.

[4] S. Joshi, P. Lorenzen, G. Gerig, and E. Bullitt, "Structural and radiometric asymmetry in brain images," in *Med. Image Anal.*, vol. 7, no. 2, 2003, pp. 155 –170.

[5] I. Bankman, *Handbook of Medical Image Processing and Analysis.* Academic Press, 2008.

[6] B. J. Gibb, *The Rough Guide to the brain.* London: Rough Guides Ltd, 2007.

[7] "Brain lobes." [Online]. Available: http://www.umm.edu/imagepages/9549.htm

[8] M. Carpenter and J. Sutin, *Human Neuroanatomy.* Baltimore: Williams and Wilkins, 1983.

[9] Wikipedia(a), "Cartesian coordinate system." [Online]. Available: http://en.wikipedia.org/wiki/Cartesian_coordinate_system

[10] Wikipedia(b), "Human anatomy planes." [Online]. Available: http://en.wikipedia.org/wiki/File:Human_anatomy_planes.svg

[11] L. Nyul, J. Udupa, and X. Zhang, "New variants of a method of mri scale standardization," in *IEEE Transactions on Medical Imaging*, vol. 19, no. 2, feb. 2000, pp. 143 –150.

[12] D. Unay, A. Ekin, and R. Jasinschi, "Medical image search and retrieval using local binary patterns and klt feature points," in *15th IEEE Int. Conf. on Image Processing*, oct. 2008, pp. 997 –1000.

[13] MRIcro. [Online]. Available: http://www.mccauslandcenter.sc.edu/mricro/mricro/mricro.html

[14] ITK-SNAP. [Online]. Available: http://www.itksnap.org/pmwiki/pmwiki.php

[15] RView. [Online]. Available: http://rview.colin-studholme.net/

[16] T. Brigham and W. Hospital, "3dslicer." [Online]. Available: http://www.slicer.org/

[17] R. A. Morey, C. M. Petty, Y. Xu, Pannu, Wagner, D. V. Lewis, K. S. Labar, M. Styner, and G. Mccarthy, "A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes," in *NeuroImage*, vol. 45, apr 2009, pp. 855–866.

[18] J. Lotjonen, J. Koikkalainen, L. Thurfjell, and D. Rueckert, "Atlas-based registration parameters in segmenting sub-cortical regions from brain mri-images," in *IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*, july 2009, pp. 21 –24.

[19] A. R. Khan, M. K. Chung, and M. F. Beg, "Robust atlas-based brain segmentation using multi-structure confidence-weighted registration," in *Proceedings of the 12th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention: Part II*, 2009, pp. 549–557.

[20] F. v. d. Lijn, "Automated atlas-based segmentation of brain structures in mr images application to a population-based imaging study," Ph.D. dissertation, Erasmus Universiteit, Rotterdam, 2010.

[21] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," in *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, jul 2002, pp. 971 –987.

[22] J. Shi and C. Tomasi, "Good features to track," in *IEEE Comput. Society Conf. on Comput. Vision and Pattern Recognition*, jun 1994, pp. 593 –600.

[23] A. Ekin, "Fature-bsed brain mid-sagittal plane detection by ransac," in *14th European Signal Processing Conference*, 2006.

[24] B. Ardekani, J. Kershaw, M. Braun, and I. Kanuo, "Automatic detection of the mid-sagittal plane in 3-d brain images," in *IEEE Trans. on Medical Imaging*, vol. 16, no. 6, dec. 1997, pp. 947 –952.

[25] Y. Liu, R. Collins, and W. Rothfus, "Robust midsagittal plane extraction from normal and pathological 3-d neuroradiology images," in *IEEE Trans. on Medical Imaging*, vol. 20, no. 3, march 2001, pp. 175 –192.

[26] S. Prima, S. Ourselin, and N. Ayache, "Computation of the mid-sagittal plane in 3-d brain images," in *IEEE Trans. on Medical Imaging*, vol. 21, no. 2, feb. 2002, pp. 122 –138.

[27] F. Kruggel and D. Yves, "Alignment of magnetic-resonance brain datasets with the stereotactical coordinate system," in *Medical Image Analysis*, vol. 3, no. 2, 1999, pp. 175–185.

[28] B. Mazoyer, "Jean talairach (1911-2007): A life in stereotaxy," in *Human Brain Mapping*, vol. 29, no. 2. Wiley Subscription Services, Inc., A Wiley Company, 2008, pp. 250–252.

[29] Q. Hu, G. Qian, and W. Nowinski, "Fast, accurate, and automatic extraction of the modified talairach cortical landmarks from magnetic resonance images," in *Magnetic*

Resonance in Medicine, vol. 53, no. 4. Wiley Subscription Services, Inc., A Wiley Company, 2005, pp. 970–976.

[30] W. L. Nowinski, "Modified talairach landmarks," in *Acta Neurochirurgica*, vol. 143. Springer Wien, 2001, pp. 1045–1057.

[31] K. B. Prakash, Q. Hu, A. Aziz, and W. L. Nowinski, "Rapid and automatic localization of the anterior and posterior commissure point landmarks in MR volumetric neuroimages," in *Academic Radiology*, vol. 13, no. 1, 2006, pp. 36–54.

[32] Y. Fu, W. Gao, X. Chen, M. Zhu, W. Shen, and S. Wang, "Automatic identification of the reference system based on the fourth ventricular landmarks in t1-weighted mr images," in *Academic Radiology*, vol. 17, no. 1, 2010, pp. 67–74.

[33] B. Ardekani and A. Bachman, "Model-based automatic detection of the anterior and posterior commissures on mri scans," in *NeuroImage*, vol. 46, no. 3, 2009, pp. 677–682.

[34] C. Don, T. Megan, and J. Gregory, "Automatic localization of the anterior and posterior commissures in mri brain images using artificial neural networks," in *Int. Soc. Mag. Reson. Med. Recent Advances in Image Analysis: Applications.*, vol. 19, May 2011, p. 2602.

[35] L. Juan and O. Gwun, "A comparison of sift, pca-sift and surf," in *Int. Journal of Image Processing*, vol. 3, no. 4, 2009, pp. 143–152.

[36] P. V. Lukashevich, B. A. Zalesky, and S. V. Ablameyko, "Medical image registration based on surf detector," in *Pattern Recognit. Image Anal.*, vol. 21, no. 3, sep. 2011, pp. 519–521.

[37] E. Nagham, F. Abou-Chadi, and S. Kishk, "Wavelet-based image registration techniques: A study of performance," in *Int. Journal of Comput. Science and Network Security*, vol. 1, no. 3, 2011, pp. 519–521.

[38] Z. Li-jia, Z. Shao-min, Z. Da-zhe, Z. Hong, and L. Shu-kuan, "Medical image retrieval using sift feature," in *2nd Int. Congress on Image and Signal Processing. CISP.*, oct. 2009, pp. 1–4.

[39] D. Lowe, "Object recognition from local scale-invariant features," in *Proc. of the 7th IEEE Int. Conf. on Comput. Vision*, vol. 2, 1999, pp. 1150–1157.

[40] Y. Ke and R. Sukthankar, "Pca-sift: a more distinctive representation for local image descriptors," in *Proc. of the IEEE Comput. Society Conf. on Comput. Vision and Pattern Recognition*, vol. 2, 2004, pp. 506–513.

[41] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, 2005, pp. 1615–1630.

[42] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking." in *IEEE Int. Conf. on Comput. Vision*, vol. 2, 2005, pp. 1508–1511.

[43] H. Bay, A. Ess, T. Tuytelaars, and V. Luc, "Speeded-up robust features (surf)," in *Comput. Vision and Image Understanding*, vol. 110, no. 3, 2008, pp. 346–359.

[44] D. Lowe, "Distinctive image features from scale-invariant keypoints," in *Int. Journal of Comput. Vision*, vol. 60, no. 2, 2004, pp. 91–110.

[45] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *CVPR*, vol. 1, 2001, pp. 511–518.

[46] T. Lindeberg, "Discrete scale-space theory and the scale-space primal sketch," Ph.D. dissertation, KTH Stockholm, 1991.

[47] M. Brown, R. Szeliski, and S. Winder, "Multi-image matching using multi-scale oriented patches," in *IEEE Comput. Society Conf. on Comput. Vision and Pattern Recognition*, vol. 1, june 2005, pp. 510–517.

[48] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, may 1997, pp. 530 –535.

[49] L. Robert, *OpenCV 2 Computer Vision Application Programming Cookbook*. Birmingham,UK: Packt Publishing Ltd, 2011.

[50] OASIS, "The open access series of imaging studies." [Online]. Available: http://www.oasis-brains.org

[51] IXI, "Information extraction from images." [Online]. Available: http://biomedic.doc.ic.ac.uk/brain-development/index.php?n=Main.Datasets

[52] ITK, "Insight segmentation and registration toolkit." [Online]. Available: http://www.itk.org/