UNIVERSITÄT ZU LÜBECK

From the Institute for Signal Processing of the University of Lübeck,Germany
Director: Prof. Dr.-Ing. Alfred Mertins

Computer-Aided Detection of Parkinson's Disease Using Transcranial Sonography

Dissertation
for Fulfillment of Requirements
for the Doctoral Degree
of the University of Lübeck
from the Departments Information/Technology

submitted by
Lei Chen
from Chifeng in China

Lübeck 2013

UNIVERSITÄT ZU LÜBECK
INSTITUTE FOR SIGNAL PROCESSING

http://www.isip.uni-luebeck.de/

First referee: Prof. Dr.-Ing. Alfred Mertins

Second referee: Prof. Dr. rer. nat. Amir M. Mamlouk

Chairman: Prof. Dr. rer. nat. Thorsten M. Buzug

Date of oral examination: September 26, 2013

Approved for printing, Lübeck: November 20, 2013

ABSTRACT

Medical imaging is a vitally necessary tool of healthcare in modern medicine. Machine leaning plays an essential role in this field, such as Computer-Aided Diagnosis (CAD). Parkinson's disease is one of the most common neurological diseases. The primary symptoms of PD result from the loss of the nerve cells that secrete dopamine in the region of the substantia nigra (SN). Although PD is currently regarded as incurable, the neurons of the SN can be sheltered by neuroprotective drugs when used in the early stages. Therefore early diagnosis of PD is of great importance, and that means preclinical diagnosis before the first parkinsonian motor symptoms occur. In 1995, Becker et al. first used transcranial sonography (TCS) to visualize the midbrain region and found an enlarged area of the SN (SN hyperechogenicity) in PD patients compared with controls. At the early stages, TCS is more suitable for the diagnosis of PD than other medical modalities, such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). In this thesis, a brief history of TCS applied to the PD diagnosis will be introduced. The limitations of the TCS method that affect the diagnosis of PD includes the accessibility of the SN in subjects, the dependence of image quality on the experience of the sonographer, the variation of measurements in different ultrasound systems and different laboratories, and the standardized determination approach of the extent of hyperechogenicity. The goal is to apply the image analysis methods onto the TCS images to detect the pattern of PD and assist the physician during the diagnostic procedure.

The thesis combines image analysis techniques with prior knowledge from experts and anatomy of brain. The medical image analysis composes of four sequential stages which are image acquisition, image enhancement, image segmentation, and image quantification. In each stage, we design and implement image processing techniques that compose the CAD based on TCS images. Specifically, this system includes a segmentation approach for the area of interest (ROI) extraction, the feature extraction methods for TCS image classification, a ROI detection for SN hyperechogenicity, and the feature selection methods for better performance of the classifiers. In this thesis, we collect and analyze TCS images that were obtained from two ultrasound machines, Philips SONOS 5500 and Siemens Sonoline Antares.

Regarding investigator dependence, a semi-automatic segmentation algorithm is applied to extract the regions of SN in the TCS images. The main content is

to design different feature extraction methods that can be developed to describe other distinct information contained in the images. These features aim at separating images of individuals that are genotypically or phenotypically different. A multiple feature extraction algorithm is proposed that computes statistical features, geometrical features, and texture features from ROIs. Afterwards, feature selection methods are used to find the best feature subset that can achieve the best classification rate. Furthermore, a robust feature extraction algorithm is developed by using a rotation-invariant Gabor filter and compute robust features based on the normalized histogram.

In this thesis, the invariant scattering convolution networks is first applied on TCS images. In order to use the scattering coefficients based on the data with small training set, the feature dimensionality reduction methods are designed and implemented. Combining with the feature selection method, the selected scattering coefficients achieved even better performance for the TCS image classification. Moreover, a classification method with LDA is proposed that instead of the PCA used in the original work to reduce the computation time, while keeping or improving the accuracy. A shape-adapted blob detection algorithm is presented to estimate the hyperechogenicity of SN in TCS images. This blob detection method can supply the visible detection results to the physician instead of the feature extraction results mentioned in the previous works. The locations of all suspected areas are positioned by a scale invariant blob detection method, and then the ROIs are estimated by using the watershed segmentation algorithm and a shape-adapted interest area detector.

Moreover, a sequence analysis method is introduced that based on the recorded images during the acquirement of TCS examination. Considering the identification of the scan plane to be investigator independent, the obtained sequence is registered and visualized in 3D space. The doctor segmentations of the midbrain is then used to segment the volume of mesencephalic stem. As a result, the better diagnosis can be made with the help from 3D visualization of SN region instead of one single 2D TCS image.

Acknowledgments

For this doctoral dissertation I have gained a lot of support and help from my dear supervisors, my colleagues, and my family.

First and foremost, I would like to express my heartfelt gratitude to my supervisor Prof. Dr.-Ing. Alfred Mertins who gave me the opportunity to work on this thesis in his research group. I appreciate him for guiding me to the area of pattern recognition and for giving me the freedom to follow my research interest. I also appreciate his academic guidance, financial support, and theoretical advice for developing this thesis.

I appreciate Prof. Ulrich Hofmann who offered me this research position and supported me involving the research environment in Luebeck University. Acknowledgment should be shown to Prof. Med. Güter Seidel for introducing me to the medical question, my research topic. Especially I like to thank my second supervisor Dr. Med. Johann Hagenah, he prepared a large data set for us meanwhile having to cope with an everexpanding workload.

This research position is supported by the Graduate School, I am grateful for the management team, Dr. Markus Finke, Katja Dau, Philipp Jauer, and Olga Schachmatova. Particularly, I appreciate Chaoqun Jiang for helping me to handle most of documentary works in Luebeck University. I also thank Katja Dau and Dr. Kerstin Lütke-Buzug for helping my family and offering my wife a part time job.

Yijing Xie and I shared an office all four years and it is my pleasure. Both study and personal life, she really helped me a lot, many thanks to her. Especially, I am indebted to Dr.-Ing Christian Kier who introduced me to Institute for Signal Processing and supported me starting the research for this thesis. I would like to thank my colleague Dr.-Ing Alexandru Condurache for his attending my progress meeting and practical suggestions for my study. I also appreciate Dr.-Ing Radoslaw Mazur for his material and advice regarding to my thesis writting. My grateful thanks to Dr.-Ing Florian Müller, Jan Ole Jungmann, Dierck Matern, and Marco Maaß for the useful discussions during my talks. I also have enjoyed all the chats both in working and relax time. This thesis requires supports from our technician, Thomas Schnelle, for maintaining the computer, and from our secretary, Frau Christiane Ehlers, for helping me to handle the documentary works. Thanks a lot. In addition, I thank Julia Krüger for the semi-automatic segmentation software and thank Arkan Al-Zubaidi for his the research work and one publication.

Finally I would like to express my gratitude to my beloved parents, my family for great confidence in me and for their loving considerations. I also owe a special debt of gratitude to my wife, Dandan Li, for her alwalys helping and supporting me without a word of complaint. I wish to express my love to her for sacrificing her career to taking care of our daughter, Sirui Chen. Thanks her lovely smile that would release all my pressure and fatigue immediately. My gratitude also extends to my parents-in-law and my three old sisters for their unfailing encouragement and support.

Contents

Chapter 1

Introduction

In 1817, James Parkinson first described the clinical syndrome of Parkinson's disease (PD). At the present time around a quarter of all patients diagnosed by neurologists have some other pathology at postmortem [1]. Therefore neurologists 'cannot make an accurate purely clinical definition' of PD [1]. The pathologists found that Lewy bodies are always present in some surviving neurons in the substantia nigra (SN) in all cases of PD[2], that is the pathological feature of PD for the clinician. But, the classical pathological features have been found at postmortem from some cases, yet they have had no any symptoms in life [1]. With a compromise, the definition of PD is 'a condition where typical clinical features are present in life and a particular pathology is found at postmortem' [1]. Parkinsonism is defined as the clinical features which are characteristic of PD. The cardinal features of PD such as Tremor at rest, Rigidity, Akinesia and Postural instability [3] can be used for the diagnosis during the ordinary clinical examination. The primary symptoms of PD result from the loss of the nerve cells that secrete dopamine in the region of the SN [4]. In 1995, Becker et al. first used Transcranial sonography (TCS) to visualize the midbrain region and found an enlarged area of the SN (SN hyperechogenicity) in PD patients compared with controls. Regarding the PD diagnosis by other medical modalities, such as Computed tomography (CT) and Magnetic Resonance Imaging (MRI), TCS is more suitable for the diagnosis of PD at the early stages [5]. In this chapter, a brief history of TCS applying on the earlier PD diagnosis will be introduced. Our goal is to apply the image analysis methods onto the TCS images to detect the pattern of PD and assist the physician during the diagnostic procedure.

1.1 Parkinson's Disease

Parkinson's disease is one of the most common neurological diseases with a prevalence of $160/100,000$ in Western Europe rising to around 4% of the population over eighty [6]. The symptoms of PD occur when neurons of the substantia nigra die or become impaired, a schematic SN region in mesencephalon is shown in Figure 1.1. Normally, these cells are responsible for the production of a chemical messenger called dopamine, which transmits signals within the brain to produce smooth physical movements. When these neurons cease to produce dopamine, the communication between the brain and muscles weakens. As a result, the brain becomes unable to control the muscle movement [7]. Most cases of PD occur in people without apparent family history of the disorder are classifed as sporadic. These cases are probably related to an interaction of genetic and environmental factors, although the reason to cause sporadic cases remains unclear [8]. Approximately 15% of PD patients have an apparent history of this disorder in their family. The genes related to the familial cases of PD are mutations in the LRRK2, PARK2, PARK7, PINK1, SNCA, or mutations in other genes that have not been identifed. Mutations in some of these genes may also play a role in the sporadic cases. The genetic causes of PD are not fully understood, and the influence of the risk of developing the disorder by the genetic changes is still under investigation. The protein deposits called Lewy bodies are found in dead or dying nerve cells in SN. However, the role of the lewy bodies whether they are response to kill nerve cells or part of the cells' response to the disorder, is still unclear so far.

In terms of their origin, parkinsonian disorders can be divided into four types: primary (idiopathic) parkinsonism, secondary parkinsonism, heredodegenerative parkinsonism and multiple system degeneration [3]. The Parkinson's disease, as the most common form of parkionsonism, is usually called idiopathic parkinsonism or primary parkinsonism to differentiate it from other forms of parkinsonism. According to the cause, familial Parkinson's disease and sporadic Parkinson's disease are used to name genetic and the idiopathic PD [6], respectively. Regarding the disease progress, PD can be divided into three stages [9]:

- Stage I, a disease-free state. In this stage the risk factors are present but the nigral neuronal cell loss is still under the normal age-related decline.

- Stage II, the early stage or pre-diagnostic phase of the disease. In this phase, the loss of neuronal cell in SN exceeds the normal age-related decline. But no or only mild symptoms, such as soft motor signs (arm

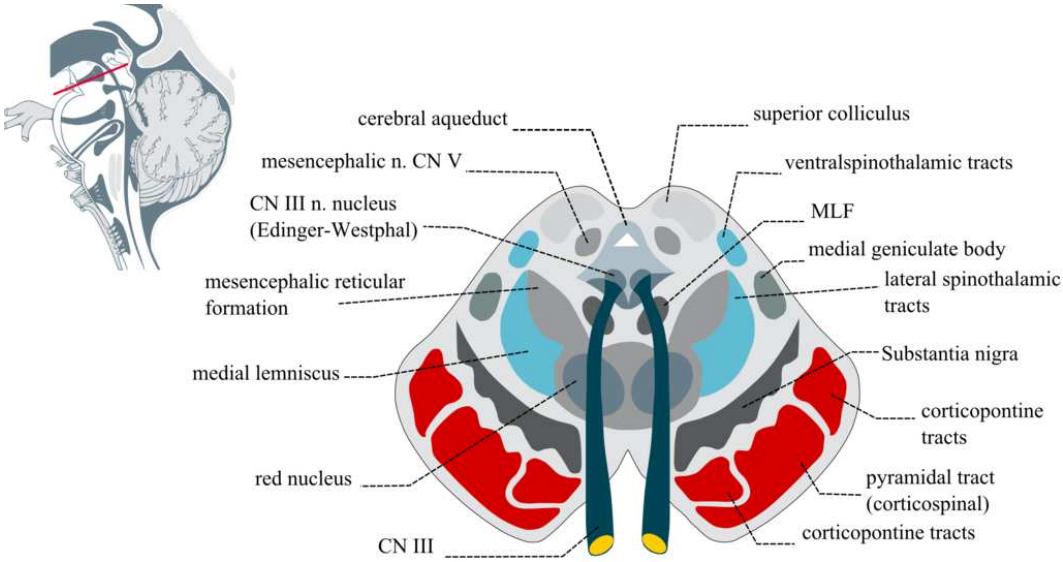


Figure 1.1: Schematic midbrain with substantia nigra, the section through superior colliculus showing substantia nigra. Reuse with permission of Wikimedia Commons, author is Madhero88.

swinging, changes in hand-writing) and non-motor signs (depression), can be detectable.

- Stage III, the motor symptoms appear and the clinical diagnosis is accepted according to the widely accepted criteria. Where 60% of dopaminergic neurons have damaged and the striatal dopamine content is deseased by 80% or even more.

Although PD is currently regarded as incurable, the neurons of the SN can be sheltered by neuroprotective drugs when used in the early stages and the symptoms can be alleviated [10]. Therefore early diagnosis of PD is of great importance and that means preclinical diagnosis of predisposed individuals before the first parkinsonian motor symptoms occur [11]. Due to the well know difficulty in diagnosing PD, one popular clinical criteria set was devised by the United Kingdom Parkinson's Disease Society Brain Bank. The first step is to diagnose a parkinsonian syndrome, such as muscular rigidity or rest tremor. The second step is to fill in a checklist of exclusion criteria such as a history of presence of signs. The third step is to find prospective criteria such as levodopa response for many years. The patient can be clinically diagnosed as 'definite Parkinson's disease' when more than three supportive criteria are present [1]. Alternatively, the diagnosis of PD includes the medical history and a neurological examination [3]. The histologic demonstration of intraneruonal Lewy bodies in the midbrain

is usually considered proof for the definitive diagnosis of idiopathic PD [12], but such demonstration on autopsy is clearly impractical. However, patients of PD typically show only age-related changes but no distinctive imaging abnormalities [13]. Conventional T1- and T2-weighted MRI shows normal brain structure in idiopathic PD [12], only patients with advancing PD T2-weighted images show increased signal that was found positively correlated to the deposition of iron in the SN [14, 15]. The diffusion-weighted MRI can play a valuable role in discriminating atypical parkinsonian syndromes from typical PD as shown in Figure 1.2, the water-proton apparent diffusion coefficients raised in the putamen in most patients with multiple-system atrophy (MSA) but are normal in PD, although the clinical diagnostic uncertainty is still present [12].

Furthermore, the function of dopamine in the brain can be measured by the Positron Emission Tomography (PET) with tracer 18-fluorodopa (FDOPA) and the Single-Photon Emission Computed Tomography (SPECT) with tracers ^{123}I-FP-CIT (DaTSCAN; GE Health care) [12, 16]. An unaffected gene LRRK2 mutation carrier with PD shows low dopamine transporter binding and a mutation carrier with the inherited form of PD is most significantly in these three subjects as shown in Figure 1.3. The demonstrated pattern of reduced dopaminergic activity with PET and SPECT can aid physicians in the diagnosis of PD. The diagnostic accuracy of SPECT in parkinsonian patients was firstly studied in University of Maastricht, the Netherlands [17]. Especially they investigated the accuracy of transcranial duplex scanning (TCD) in the diagnosis of PD, and the combination of both techniques was assessed. In addition, a new medical imaging technique, high-resolution MR imaging at $7T$ is used to identify the anatomy of midbrain dopamine regions [18]. They scanned the anatomic structure of the midbrain by using gradient- and spin-echo (GRASE) MR imaging and 3D gradient-echo sequences. GRASE is a $T2$ and $T2^*$ weighted multi-shot imaging sequence [19]. GRASE scan acquires data in 2 dimensions, the gradient-echo technique with fully refocused transverse magnetization (balanced FFE) [20] acquires data in 3D. The results of both GRASE and FFE scans revealed visible contrast in the midbrain regions. Especially, 'the FFE scan also displayed distinct contrast between subregions of the SN showing sensitivity to iron-related magnetic susceptibilities' as shown in the Figure 2 in [18].

1.2 Transcranial Sonography

In 1995, Becker et al. first described and used Transcranial sonography (TCS) in a clinical study between a small group of Parkinson's disease (PD) patients and

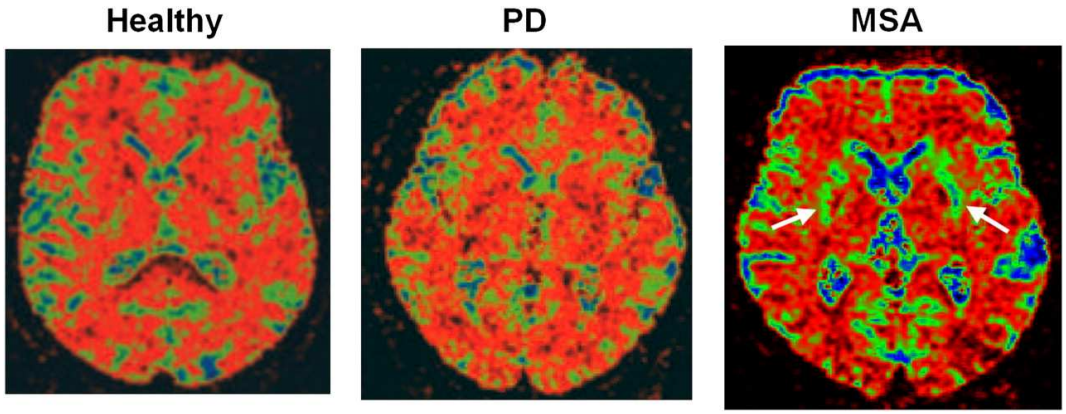


Figure 1.2: Color-coded diffusion weighted MRI images. In striatum of PD patient the apparent diffusion coefficient is normal but raised in (multiple-system atrophy) MSA. Reuse with permission of [21].

healthy controls [22]. Since the 1980s, transcranial color-coded duplex sonography (TCCS) has been applied for diagnostic ultrasonography in the central nervous system. Compared to conventional transcranial pulsed-wave Doppler (TCD), TCCS has more decisive advantages [23]. During using TCCS and TCD in clinical application, adult transcranial B-mode sonography, a full name of TCS, has evolved as an extension of the experiences of sonography [23]. TCS is permitted to assess the ventricular system at that moment. In general, TCS is performed with an ultrasound machine attached with a phased-array using pulse-echo technique, which provides a two-dimensional image of the butterfly-shaped midbrain [24]. The schematic illustration of the scanning plane at the midbrain and the corresponding MRI and TCS images are shown in Figure 1.4. In TCS images, an enlarged area with significantly increased echogenicity of the substantia nigra (SN hyperechogenicity) was found in PD patients compared with controls. This finding the SN hyperechogenicity in the TCS images of PD was confirmed by another independent group in 2002 [25]. Although CT and MRI brain scans of PD appear normal, the SN shows a distinct hyperechogenic pattern on TCS images in about 90% of PD patients [9]. The reason why the signal intensity of SN is increased on TCS images of the PD patients is suggested to be an increased iron concentration in the SN, causing oxidative stress and neuronal cell damage [9]. At early stages it is possible to determine the formation of idiopathic PD as well as monogenic forms of parkinsonism by means of TCS [26]. Furthermore, the SN hyperechogenicity was found to associate with a significant reduction of 18-fluorodopa (FDOPA) uptake in the striatum measured with PET [5]. These studies show that the SN hyperechogenicity is

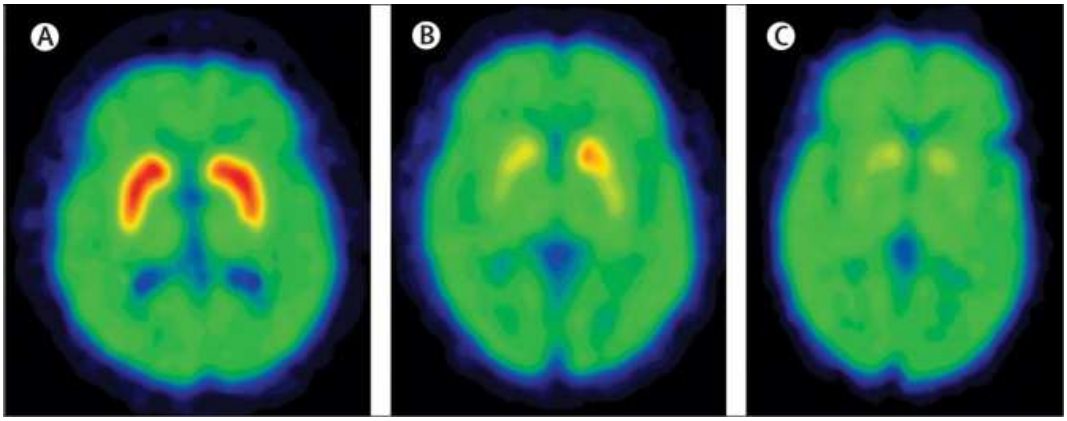


Figure 1.3: Dopamine transporter binding on PET imaging in a healthy individual (A), a clinically unaffected LRRK8 mutation carrier(B), and a patient with LRRK8-related Parkinson's disease (C). 'The graded and asymmetrical reduction in dopamine transporter binding, with the greatest amount of binding in the healthy individual and the least in the patient with LRRK8-related Parkinson's disease.' Reuse with permission of [16].

a valuable marker for PD diagnosis, especially for early diagnosis [27]. However, the image resolution of TCS is limited because of the low frequency of the transduser. The image quality mainly dependents on the acoustic bone window of the individual. In addition, the image properties, such as the brightness and the scale, might be affected by the experience of the examiner because of the different settings of ultrasound machine. With increased use, a standardized procedure is required [13] that including rating scale of SN echogenicity.

1.2.1 Scanning Protocol and Clinical setting

Compared to the other medical imaging modalities, TCS is a low-cost, noninvasive and mobile method which can be performed with unlimited repetition. This method was facilitated by the technical improvement in B-mode ultrasound machine with phased-array probe [13]. The examination is performed with positioning of the probe through the posterior temporal bone window (less than 2 mm) that is the most commonly used. The low frequency range of ultrasound wave is set to 1.6 to 2.5 MHz, because the high frequency (> 4 MHz) ultrasound wave cannot penetrate to the deep brain through the bone window [29]. As a result, the spatial resolution of TCS images is lower than for other scanning of

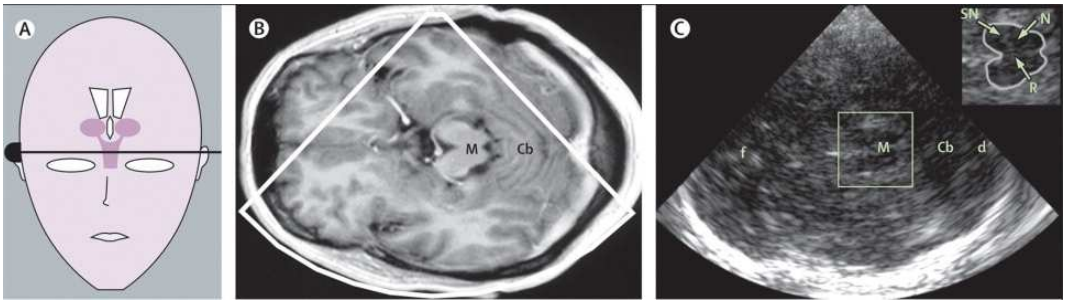
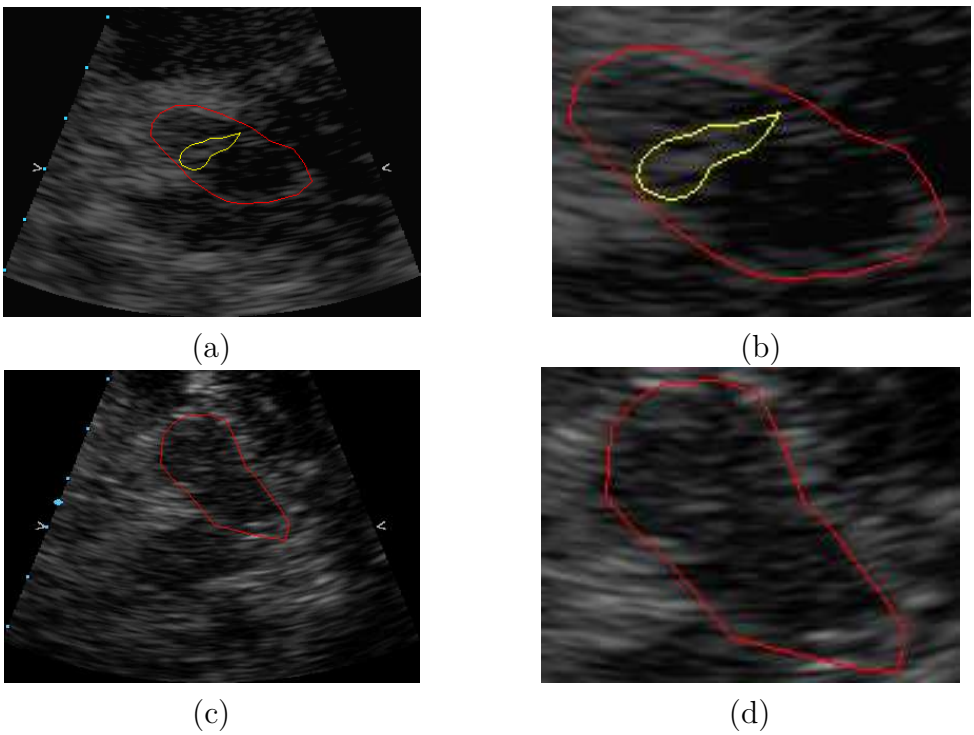


Figure 1.4: The illustration of the axial scanning plane and the corresponding MRI and TCS images. (A) Schematic illustration of the axial scanning plane at the level of the midbrain. (B) MRI image of axial section at midbrain level. (C) The corresponding TCS image at midbrain. The magnified square area at the upper right corner indicates the mesencephalon and SN structures. M is mesencephalon, Cb is cerebellum, d is dorsalare, N and R are red nucleus and raphe, respectively. Reuse with permission of [28].

the soft tissue: the axial resolution is approximately 0.7 mm and the lateral resolution varies between 2.2 and 3.8 mm, depending on the ultrasonic beam [30]. The parameter of dynamic range is often set to 50-55 db with a penetration depth of 14-16 cm [29]. The investigation by TCS is conducted according to a standardized protocol in distinct scanning planes such as mesencephalic plane, third ventricular plane and cella media plane, by certain landmark structures (Figures 1-5 in [29]). The SN hyperechogenicity is assessed in the butterfly-shaped mesencephalon on the mesencephalic scanning plane, the corresponding MRI and TCS images of each plane can be seen in [24]. Considering the decreased image quality, signal-to-noise ratio with increasing insonation depth, only the ipsilateral *half of mesencephalon* (HoM) which is close to the probe is examined by the physician. Therefore, in a routine clinical examination two TCS images from left and right side of the brain are acquired per individual subject for the diagnosis. The regions of HoM (ROIs) and SN are subsequently manually-marked by the physicians as shown in Figure 1.5. Compared with the healthy controls, the size of the hyperechogenic SN area is relatively large from the TCS image of PD patients.

The size measurement of SN hyperechogenicity is performed on individual scan after manually marking the outer circumference of SN echogenic area. The size of SN echogenic < 0.2 cm^2 is classified as normal, the area size of 0.25 cm^2 and above as markedly hyperechogenic, and size in-between as moderately hyperechogenic [22, 25].



(a) (b)

(c) (d)

Figure 1.5: TCS images marked by physicians, Philips SONOS 5500. The images in the upper and bottom row that were collected from a PD patient and a healthy control subject, respectively. (a) and (c): The butterfly-shaped midbrain images on the mesencephalic plane; (b) and (d): The region of the ipsilateral mesencephalon wing that is close to the probe. The red marker indicates the upper half of mesencephalon. Yellow markers show the SN area as a bright spot.

1.2.2 Main Factors of TCS

The limitations of the TCS method that affect the diagnosis of PD include the accessibility of the SN in subjects, the dependence of image quality on the experienced sonographer, the variation of measurements in different ultrasound systems and different laboratories, and the standardized determination approach of the extent of hyperechogenicity. First, the propagation of the ultrasound waves through the temporal bone window are affected by attenuation and refraction of skull bone. Therefore, the clear TCS images are difficult to obtain through the acoustic bone window because the thickness is too small (around 2 mm), especially a high rate of recording failure of SN in aged female subjects [31] in Japan. Second, only the ultrasound waves with low frequency (1-3 MHz) can

penetrate through the bone window to the deep brain for obtaining the image of midbrain. As a result, the lower frequency, which corresponds to limited resolution of the ultrasound image, affects the accuracy of the image analysis. Third, the TCS images are obtained by a trained sonographer who can follow the standardized scanning procedure, but the probe is positioned to the head of the subject manually and the identification of the scanning planes are also investigator dependent. In addition, the lateral resolution depends on different widths of the ultrasound beam that differs between ultrasound systems [29]. The variation of measurement in different ultrasound machines and different laboratories needs to be taken into account. At last, although the SN hyperechogenicity is graded according to the semiquantitative visual rating scale [24, 29], but both the area and the brightness of SN hyperechogenicity should be considered for the quantitative analysis [5].

1.2.3 Experimental Materials

In this study, TCS images were obtained from two ultrasound machines, Philips SONOS 5500 and Siemens Sonoline Antares by three examiners. All study subjects underwent a detailed neurological examination in the local movement disorders team at Luebeck University. The assessment includes the Unified PD Rating Scale (UPDRS) and Hoehn-Yahr stage on medication. Except subjects with positive family history, PD was defined according to the United Kingdom PD Brain Bank Criteria. The examiners performed TCS with Philips SONOS 5500 ultrasound machine (Philips Medical Systems, Best, the Netherlands) connected with a 2.0-2.5 MHz sector transducer (S4 probe; Philips). The maximum depth of the scan was set as 12 cm from the temporal bone window. The scan was performed from both sides of the brain but only the ipsilateral SN was evaluated in the axial mesencephalic plane (landmark: butterfly-shaped midbrain). When the midbrain was visible as clearly as possible, the image was magnified 2-fold (zoom in) and longitudinal loop comprising around 50 images of mesencephalon that were recorded for the next step study, the offline image analysis. Then, the investigators (physicians) selected two images (each from each side of brain) from around 100 stored images and rated these two images. The area (aSN) and/or mean brightness (bSN) of the echogenic SN were calculated manually by using a public-domain graphics software tool (Scion Image, Release 4.0.2, Scion Corporation, Frederick, MD, USA). Especially, the sonographers who acquired the TCS images were blinded to the results of the clinical interviews and the genetic status of these subjects. The investigators who analyzed the recorded images had not been involved in the sonographic examination.

The data from Philips SONOS 5500 includes two datasets from PD, Parkin mutation carriers and healthy controls (HC). These TCS images were acquired by different examiners of the same group in half a year. Dataset 1 and Dataset 2 were collected in two different periods and have the same group structure as listed in Table 1.1. Totally, the data includes 66 images from 37 PD patients (groups 1 and 3), 58 images from 33 Parkin mutation carriers (groups 1 and 2), and 46 images from 25 healthy subjects (group 4). All 74 study subjects underwent a neurological examination by physicians, and all these 134 TCS images were manually segmented by examiners during the diagnosis. As a result, the half of mesencephalon and SN regions of the TCS images were marked with the colored curves as shown in Figure 1.5, the red and yellow curves indicate the manual segmentations of the half of mesencephalon and SN, respectively.

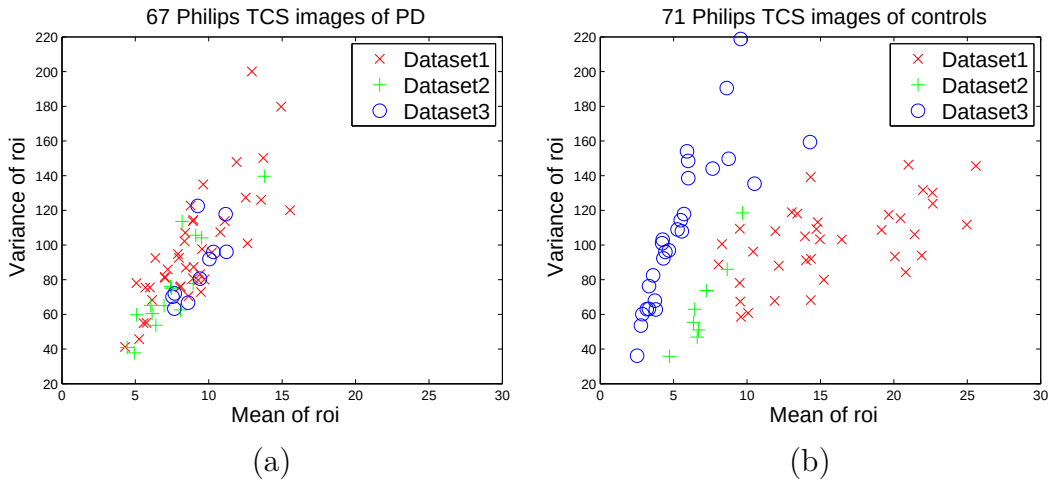
Table 1.1: The group structure of Dataset 1 and Dataset 2. The subjects in group 1 are PD patients, meanwhile, they are the Parkin mutation carriers. The subjects in group 4 are the healthy controls.

			TCS images	
Group	PD	Parkin mutation	Dataset 1	Dataset 2
*group*1	$\surd$	$\surd$	23	13
*group*2	$\times$	$\surd$	19	3
*group*3	$\surd$	$\times$	28	2
*group*4	$\times$	$\times$	38	8

In addition, the third dataset (dataset 3) includes Parkin mutational analysis in 27 subjects. There are 16 images from eight healthy controls with familial PD, 28 images from 14 healthy subjects without PD, and 10 images from five PINK mutation carriers. These subjects were screened for Parkin mutations using a comprehensive protocol mentioned above and were tested the entire coding region.

Actually, the properties of the TCS images, such as the gray values, the brightness, and the contrast, could be possibly affected by the different settings of the ultrasound machine used by different examiners. The considerable variability among different datasets was illustrated in the previous work [32]. The statistical features, the mean and variance of the region of interest (ROI) in each TCS image, were calculated and shown in Figure 1.6. The variation between each dataset can be seen from TCS images in Figure 2.1 of Chapter 2.

We also collected TCS images for the study by using a different ultrasound system, Siemens Sonoline Antares (Elegra, Siemens). A small study on TCS in



(a) (b)

Figure 1.6: The illustration about mean and variance of ROI (half of mesencephalon) of 138 TCS images. (a) 38 subjects of Parkinson's Disease. (b) 39 subjects of healthy control.

seven subjects used both Philips SONOS 5500 and Siemens Sonoline Antares ultrasound systems. This study aimed to analyze the difference between the images obtained from different ultrasound systems. Except for the same subjects study, the data collected from Siemens Sonoline Antares includes 36 subjects, 15 PD patients and 21 healthy control subjects, in total of 72 TCS images.

1.3 Methods of Parkinson's Disease Computer Aided Detection

Medical imaging is a vitally necessary tool of healthcare in modern medicine. In this field, machine leaning plays an essential role which includes Computer-Aided Diagnosis (CAD), medical image analysis, image-guided therapy [33]. The first commercial product of Computer-Aided Detection approved by the U.S. Food and Drug Administration is the system for breast imaging [34]. Besides breast imaging, computer-aided diagnosis (CADx) systems are also applied in the area of thoracic imaging, abdominal imaging, brain imaging, and body imaging. Suzuki described major technical advancements and research findings in the field of CAD and collected more 20 examples of CAD systems in his book [33]. Although CAD systems have been applied as commercial products, the study and research still continue, such as computer-aided detection of breast

cancer using ultrasound images [35]. CAD system has been developed to aim at helping physicians to evaluate medical images and detect lesions, in the meantime to increase the detection and diagnosis accuracy and save labor [35].

The aim of this thesis is to employ computers to assist neurologists in the diagnosis of PD with TCS images. The thesis combines biomedical image analysis techniques with prior knowledge from anatomy of brain and experts. Biomedical image analysis is a highly interdisciplinary field, which is related to computer science, physics, medicine, biology, and engineering [36]. In general, biomedical image analysis is to apply image processing techniques to biological or medical problems. Biomedical image analysis composes of four sequential stages which are image acquisition, image enhancement, image segmentation, and image quantification. In each stage, we design and implement image processing techniques that compose the CADe based on ultrasound images. Specifically, with increased use of TCS during routine examination, a standard clinical setting including rating scale of SN hyperechogenicity is required. However, this technique is still based on manual evaluation by the physicians. In oder to reduce the investigator dependence, we design and develop the computer aided system for the PD detection. This system includes a segmentation approach for the area of interest (ROI) extraction, the feature extraction methods for TCS image classification, a ROI detection for SN hyperechogenicity, and the feature selection methods for better performance of the classifiers. The main structure of this system is illustrated as a diagram in Figure 1.7.

In this study, a large TCS dataset is analyzed that is relevant to clinical practice and includes the variability that is present under real conditions. A major difficulty of TCS image classification comes from the variability within TCS images and the influence of the user settings for the ultrasound machines. First, a semi-automatic segmentation method is applied on the midbrain region and the ROIs are extracted for further processing in following steps. Second, multiple features are extracted from ROIs, including statistical, geometrical, and texture features for the early PD risk assessment [37, 38]. Another challenge of the classification using Gabor filters is that the orientations and shapes of the mesencephalon vary from one PD patient to another. We have two solutions for this problem, one is the rotation-invariant Gabor filter with a robust feature extraction based on entropy [39], another is the shape normalization of the half of mesencephalon by a designed image warping technique. Furthermore, a local feature analysis method is proposed to detect the distinct pattern of SN in half of mesencephalon region with the localization and the quantity analysis of the interesting areas. This approach is based on blob detection and watershed segmentation [32] and that can indicate the suspected blobs in midbrain region

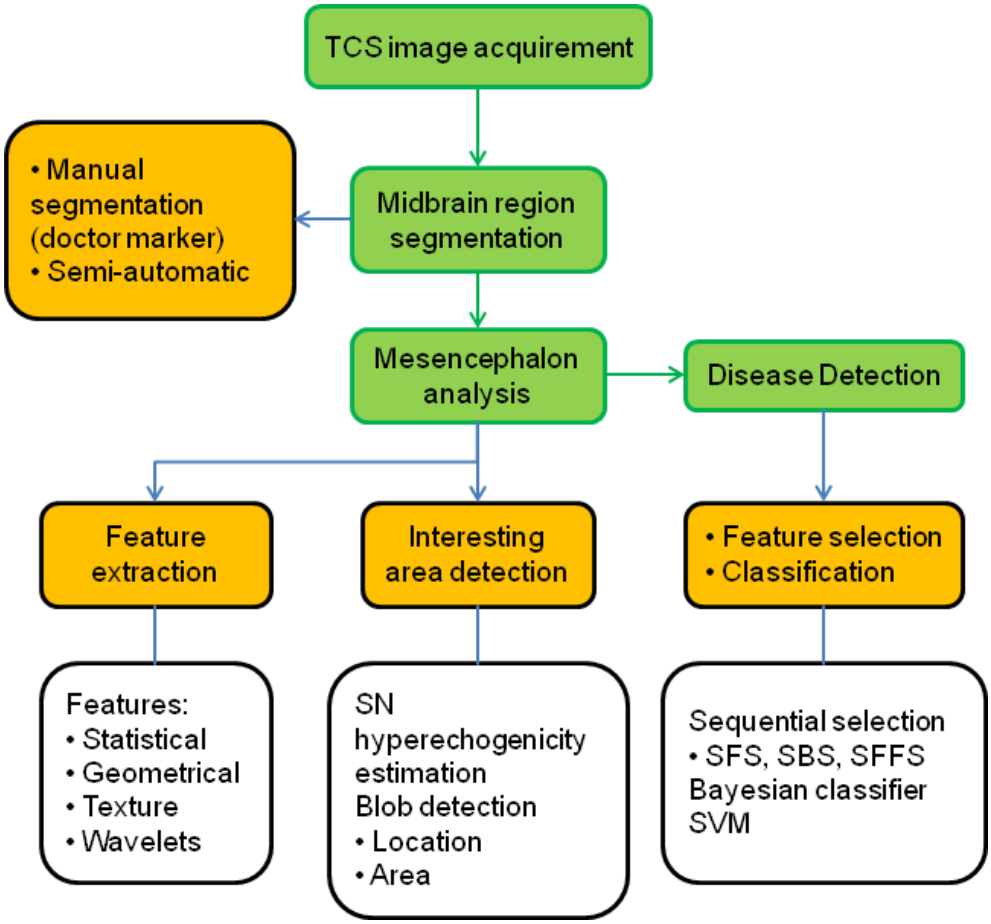


Figure 1.7: The main structure of Parkinson's Disease Computer Aided Detection system.

with a quantitative analysis.

1.3.1 Image Segmentation

The segmentation of the midbrain and SN is the crucial part for the diagnosis of PD by means of TCS. The first automatic segmentation of the midbrain and SN in 2D TCS was proposed in 2008 [40]. This method combined active contour models with a complex finite element model of midbrain anatomy. However, this method was only evaluated on ten datasets and compared with manual segmentations of the expert. The first semi-automatic midbrain segmentation from 3D TCS was implemented in 2011 [41]. But this approach did not consider

the segmentation of SN. The accurate and robust segmentation of the mesencephalon and SN from TCS images is an extremely difficult task because of the poor image quality. Hence, another automated segmentation based on the B-scan sequence of 2D TCS was proposed, which applied a complicated preprocessing for the suppression of the speckle noise to improve the active contour segmentation [42]. The first automatic SN echogenicity analysis in 3D TCS was proposed based on random-forest in 2012 [43]. In their method, the volumetric SN echogenicity detection depends on the quality of the reconstructed volume from the obtained B-scan sequences of the 2D TCS images. In this thesis, we focus on the robust image analysis method for the SN echogenicity detection from 2D TCS images. Therefore the regions of half of mesencephalon in TCS images were segmented by physicians manually and/or the applied semi-automatic segmentation method. Then the ROIs (half of mesencephalon) were analyzed for the detection of PD and the estimation of the SN hyperechogenicity.

1.3.2 Feature Extraction

Since the effective segmentation method of SN is a challenge task currently, the robust feature extraction for the TCS classification is rather more promising solution for PD detection. In our previous studies, one solution to detect PD from TCS images was to apply feature analysis on the region of the ipsilateral mesencephalon wing, which is close to the ultrasound probe. First, the moment of inertia and Hu1-moment of manually segmented half of mesencephalon were used for separating control subjects from Parkin mutation carriers [27]. Second, a hybrid feature extraction method which included statistical, geometrical, and texture features for the early PD risk assessment was proposed [37], which showed good performance of texture features (especially Gabor features). Then, a texture-analysis method that applied a bank of Gabor filters and gray-level co-occurrence matrices (GLCM) on TCS images was investigated [38]. Gabor features and GLCM texture features were combined as a feature subset with sequential forward floating selection (SFFS). The feature subset showed good results with the cross validation method.

The scattering transform is a cascade of wavelet decompositions, complex modulus operators, and local averaging. Scattering coefficients are computed with a convolution network [44], they provide much richer structure information and multi-scale texture variations [45]. We apply the invariant scattering convolution networks on TCS images and demonstrate experimental results on the classification between images from PD patients and healthy controls. However,

the scattering image representation is much larger than the original one. Therefore, the dimensionality-reduction methods on the scattering coefficient vector are investigated in this thesis. By using the scattering coefficients as feature vector, the computation time for classification is large, even if a simple classifier is used. Hence we propose to use linear discriminant analysis (LDA) instead of principle component analysis (PCA) used in the original work to reduce the feature vector for classifier while trying to keep or improve the accuracy.

Furthermore, a large dataset is analyzed with a local feature analysis method that is based on the blob detection and watershed segmentation [32]. One of the experimental results show that these local features from the detected blobs and watershed regions are largely invariant to the image normalization. Moreover, a shape-adapted interest area detector is implemented to estimate the hyperechogenicity with a large data set. This detection method is invariant to scale and affine transformation.

1.3.3 Image Sequence Analysis and Visualization

TCS is a dynamic scanning that a sonographper moves a transducer to the positions and orientations with the scanning procedure. Apparently it is difficult for a sonographer holding the transducer in a fixed position and to decide the proper images for the diagnosis. In addition, the features mentioned in above sections that were extracted from 2D TCS images cannot supply the volume information of the mesencephalon and SN. Using the current ultrasound machine, the sonographer could capture a sequence of B-scans (TCS images) from both sides of each subject. The TCS images used in this experiment were acquired with Siemens Sonoline Antares. This algorithm is designed for the analysis of TCS sequence and visualization of mesencephalon and SN region.

1.4 Outline

During the diagnosis of PD, TCS had been shown a distinct hyperechogenic pattern in images of most of PD patients. The combination of clinical characteristics and the ultrasound pattern assists in establishing the correct diagnosis of a specific movement disorder. Monogenic forms of parkinsonism may be clinically indistinguishable from PD. We and others described SN hyperechogenicity even in asymptomatic carriers of single heterozygous Parkin mutations with or without PET abnormalities. This interesting finding leads to the hypothesis

that SN ultrasound patterns may be a potential preclinical marker to detect PD susceptibility. To date, for quantitative analysis of SN hyperechogenicity, only the area of SN (aSN) but not other signal characteristics have been considered.

The original assignments of this doctoral thesis are: First, can SN hyperechogenicity serve as a preclinical marker? Second, is the ultrasound investigation useful to screen a large population for genetic and other forms of PD, thereby reducing the need for expensive genetic tests? Third, is there information other than SN hyperechogenicity in the ultrasound signal from the mesencephalic and diencephalic ultrasound images to characterize distinct forms of parkinsonism/PD?

In order to remove investigator dependence in quantifying the hyperechogenicity, we develop a semi-automatic segmentation algorithm to extract the regions of interest in the ultrasound image (SN and the surrounding mesencephalic brain stem). In addition, an image sequence analysis method has been implemented for the visualization of all the recorded images during the acquirement of TCS examination. The main content is to design different feature extraction methods that can be developed to describe other distinct information contained in the images (e.g. based on the theory of moments, regional descriptors, etc.). These features aim at separating images of individuals that are genotypically or phenotypically different.

The first chapter is the introduction about the medical background of PD and the history of TCS applying for diagnosis. The chapter 2 has given an overview of the related works on this topic, the midbrain region segmentation methods. Especially, the applied semi-automatic half of midbrain segmentation algorithm is briefly explained and the segmented results for the TCS images are compared with the manual segmentation of physicians in this thesis. In addition, the preprocessing techniques are applied before the feature extraction algorithms. The motivation of using these pre-processing methods and the details are mentioned in this chapter.

In the third chapter, a multiple feature extraction algorithm is described that compute statistical features, geometrical features, and texture features from ROIs. Afterwards, feature selection methods are used to find the best feature subset which can achieve the best classification rate. Furthermore, we develop a robust feature extraction algorithm by using a rotation-invariant Gabor filter and compute robust features based on the normalized histogram.

In the forth chapter, we first apply the invariant scattering convolution networks on TCS images. This new technique was introduced by Professor Stéphane Mallat in 2010. Based on the convolution output, the scattering coefficients, we design and develop the feature dimension reduction methods, the classification method with LDA replaces the PCA used in the original work.

In chapter 5, the hyperechogenicity of SN is estimated by a shape-adapted blob detection algorithm. The motivation is to supply the visible detection results to the physician instead of the feature extraction results mentioned in our previous works. The locations of all suspected area are positioned by a scale invariant blob detection method, and then the ROIs are estimated by using a shape-adapted interest area detector. The comparison between the estimation results and the evaluation of the doctor is given in the experiment section.

The sixth chapter considers that the identification of the scan plane is investigator independent, then the sequence obtained during the TCS examination is utilized to visualize hundreds of B-scans in 3D space by MeVisLab software. Due to unexpected movements during the acquisition procedure, the registration of each image in a sequence become very important. Therefore a local descriptor, SIFT, is used to align all images in a sequence. Moreover, the doctor marker of the midbrain is used to segment the region of mesencephalic stem in the sequence, as a result, the better performance of the visualization of SN region can be achieved.

Chapter 2

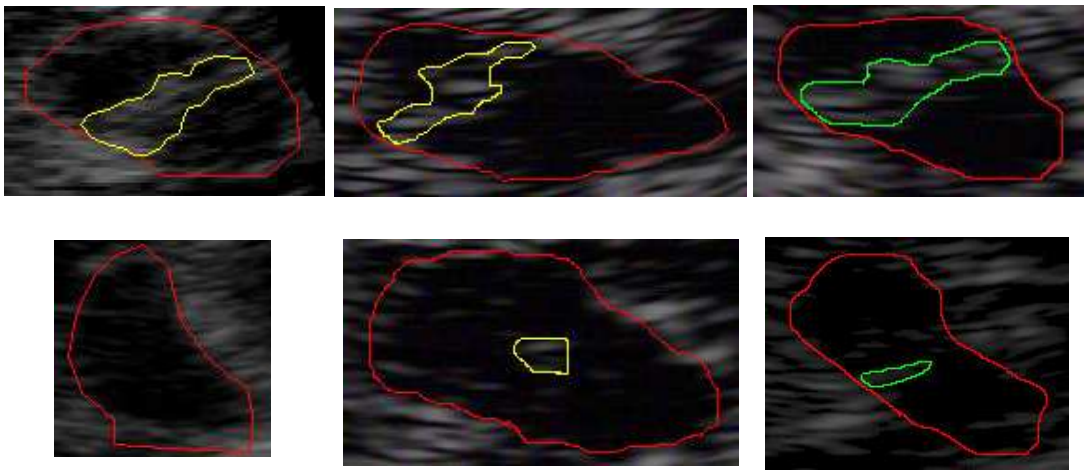
Transcranial Sonography Image Segmentation

2.1 Summary of Segmentation Methods for TCS Image

In general, the upper half of mesencephalon and the SN region are marked by physicians during a clinical examination of PD [22, 25, 9, 26]. The size of hyperechogenicity in SN area and/or the brightness of SN region are then used for the PD diagnosis. The segmentation algorithm of the midbrain and SN based on TCS images is still under investigation. In this chapter, I first review the existing segmentation methods. Then, a semi-automatic segmentation approach is applied to extract the ROIs for the further feature analysis.

2.1.1 Manual Segmentation by Physician

Based on the prior knowledge, for instance, the anatomic structure of the brain, a physician marks the area of whole or half of mesencephalon from TCS images as shown in Figure 2.1. From the view of image analysis, the mesencephalon is manually segmented. Then, several feature analysis methods are implemented [27, 37, 38, 46] using these manual segmentations from doctors. The SN hyperechogenicity is estimated by the scale-invariant blob detection method in [32]. According to this experiment, the SN area is consisted of a number of blobs in most TCS images. As a result, the SN area cannot be segmented simply just by one single curve. To date, the segmentation approach of SN area based on



(a) (b) (c)

Figure 2.1: Manually segmented TCS images from Philips SONOS 5500. The images in the first and second row are collected from PD patients, and healthy control subjects, respectively. The red marker indicates the upper half of mesencephalon. The images in each column are selected from different datasets. Yellow/green markers show the SN area as a bright spot.

one 2D TCS image is very difficult and still under investigation. Hence, only ROIs (half of mesencephalon) are considered and used in the image analysis and feature analysis algorithms for the PD detection. In this thesis, the manually segmented TCS images from Philips SONOS 5500 are used for image analysis, six of them are shown in Figure 2.1.

2.1.2 The Existing Segmentation Techniques

In order to reduce the investigator dependence, several semi-automatic algorithm were developed to segment mesencephalon or SN area [47, 40, 48, 41]. The first automatic segmentation of the midbrain and SN in 2D TCS was proposed in 2008 [40]. This method combines active contour models with a complex finite element model of midbrain anatomy, the two-component shape model [49]. This model represents the structure of the midbrain by discrete shapes as illustrated in Figure 2.2. The global model $\mathcal{T}^{(2)}$ (Figure 2.2 (c)) represents the butterfly-shaped midbrain, the local models $\mathcal{T}_i^{(1)}, i = 1, 2$, represent the stripe-like SN on each wing of mesencephalon [49]. However, the evaluation of this method was

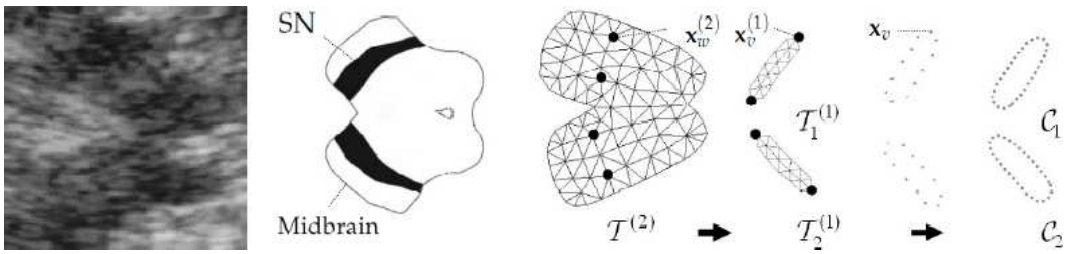
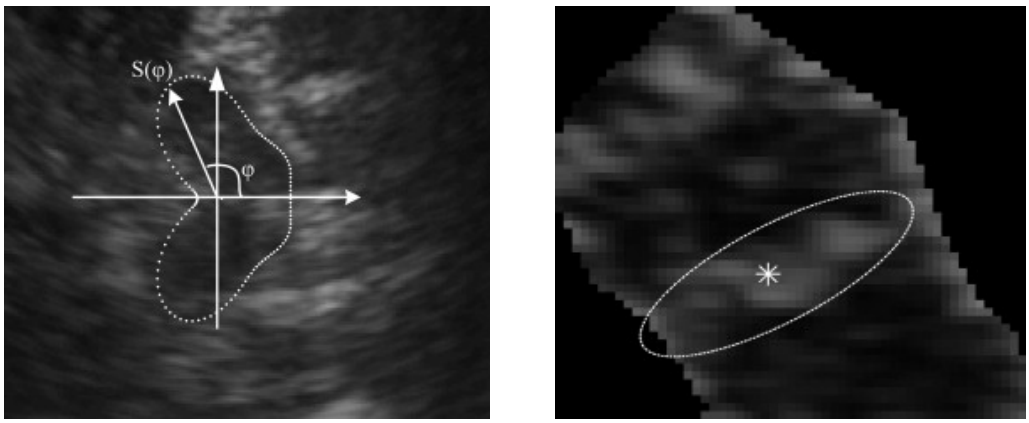


Figure 2.2: Illustration of applying the two-component shape model into the TCS images for the midbrain segmentation. (a) Midbrain region in a TCS image. (b) Schematic midbrain with stripe-like SN. (c) The corresponding two-component shape model. (d) The boundary finite element nodes of the SN-modes and the created active contours. Reuse permission of [49].

only tested on ten data sets with manual segmentations by the expert.

The accurate and robust segmentation of the mesencephalon and SN from TCS images is an extremely difficult task because of the poor image quality. Hence, another automated segmentation of 2D TCS was proposed, which applied a complicated pre-processing for the suppression of the speckle noise to improve the active contour segmentation [42]. After pre-processing of the sequence of the TCS images, a modified active contour was applied. With the assistance of the expert, all manually marked areas were averaged, and then the initial contour of the midbrain was created as shown in Figure 2.3 (a). In addition, an ellipse was selected as the initial contour for SN segmentation as shown in Figure 2.3 (b). Technically the parameters of the contour were chosen based on the prior anatomical knowledge about the SN, such as the size and the rotation.

The first semi-automatic midbrain segmentation from 3D TCS was implemented in 2011 [41]. The interesting and important step in their work is the data acquisition as sequences of 2D TCS images. They combined a medical ultrasound machine at 3 MHz with a navigation system that can record the position of the freehand probe in 3D. The scans were performed bi-laterally, as a result, the entire midbrain area can be constrained with the images obtained through the left and right temporal bone window. Based on the two sequences obtained from both sides of the brain, a 3D freehand US volume was reconstructed by using a backward-warping compounding technique [41] at resolution of 0.45 mm. The segmentation result of one 3D freehand US volume data are shown in Figure 2.4. But this approach did not consider the segmentation of SN. The first automatic



(a) (b)

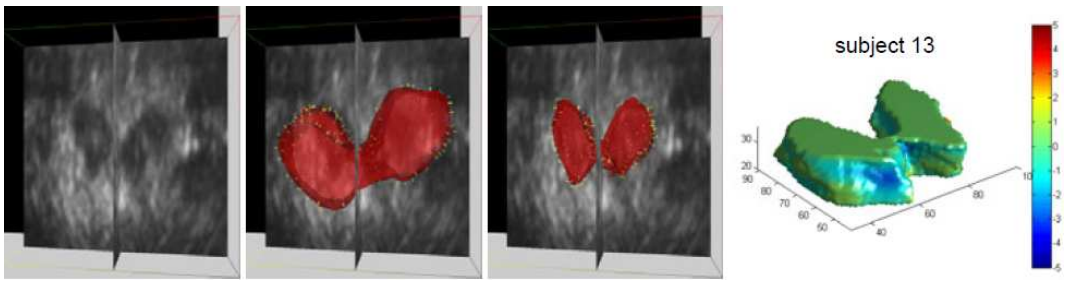
Figure 2.3: The initial contours of midbrain and SN. (a) TCS image with the approximated initial contour. (b) The defined ROI of the TCS image with a placed initial contour for the segmentation of the SN area (intensity-weighted centroid marked as *). Reuse permission of [42].

SN echogenicity analysis in 3D TCS based on random-forest was proposed in 2012 [43]. In their method, the volumetric SN echogenicity detection depends on the quality of the reconstructed volume from the obtained B-scan sequences of the 2D TCS images.

2.2 Applied Segmentation Method

Medical image interpretation is a difficult task due to the inter- and intra-personal variability existing in biologic structures [50]. A shape-based model matching algorithm, such as an active shape model (ASM), uses deformable models or atlases to match the boundaries of the object or structure in medical images. An appearance-based algorithm, such as active appearance model (AAM) [51], can represent not only the information near the landmarks but also the texture in the whole image region covered by the target object. AAMs are commonly applied to model faces [52] and they have also been used for medical image analysis [50]. In this work, the regions of half of mesencephalon in midbrain images are segmented mainly based on AAMs and the ROIs are subsequently used for feature analysis.

The idea of the segmentation is to create a 'golden' image of midbrain (anatom-



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

Figure 2.4: Illustration of the midbrain segmentation in transcranial 3D ultrasound. (a) The slice with visible midbrain. (b) and (c) Segmentation without and with data term localization, respectively. (d) The mesh surface distance map between result and ground truth. With kind permission of Springer Science+Business Media [41].

ical atlas), an initial contour, by using all TCS images labeled by experts as the training set, then match the atlas to a target image to interpret the half of mesencephalon. The advantage of AAM is that it can represent both the shape and texture variability in midbrain region in the training set. Giving such a set, a statistical model of shape variation can be generated, for details see in [53]. The labeled points on the upper half of mesencephalon in a TCS image describe the shape of half wing of the mesencephalon, which is similar to an ellipse. The shape of an AAM can be defined as a vector s, the vertex locations of the 2D triangulated mesh:

$$s = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \\ v_1 & v_2 & \cdots & v_n \end{pmatrix}. \tag{2.1}$$

We then align the sets of vectors into a common co-ordinate frame [52] and generate the model of shape variation by applying principal component analysis (PCA). Mathematically, the shape s is represented as a mean shape $\bar{s}$ with a linear combination of m shape parameter s_i:

$$s = \bar{s} + \sum_{i=1}^{m} p_i s_i, \tag{2.2}$$

where p is a set of orthogonal modes of variation [52].

The gray-level appearance model is built by warping every labeled image into the base mesh $\bar{s}$. The control points of each image are matched to the mean shape $\bar{s}$ by using a triangulation algorithm [50]. After the matching, the intensity values of the pixels are sampled from the shape-normalized image over the region covered by the base mesh $\bar{s}$. The resulting samples are normalized to reduce

the effect from the global lighting variation. Actually, obtaining the mean of the normalized data is done recursively. The details are given in [51]. The appearance of the AAM, g, is defined by the mean normalized appearance $\bar{g}$ with a linear combination of l appearance images g_i:

$$g = \bar{g} + \sum_{i=1}^{l} \lambda_i g_i, \tag{2.3}$$

where λ_i are the appearance parameters. Further details of AAMs can be found in [51, 50]. How to generate an AAM model instance was described in [54]. Two examples of the initialization of AAM on TCS images are demonstrated in Figures 2.5 (a) and (c), the finally generated contours are superimposed on the original images, shown in Figures 2.5 (b) and (d), respectively.

2.3 Preprocessing Techniques

2.3.1 Simulated Data

Considering the affections from user setting of ultrasound machine defined by different examiners, such as the brightness and contrast adjustment applied to the original TCS images, we applied four methods to TCS images to simulate the settings of the examiner. The first method is to rescale the TCS image to the range $[0, 255]$, for each image the intensity values are rescaled as given by

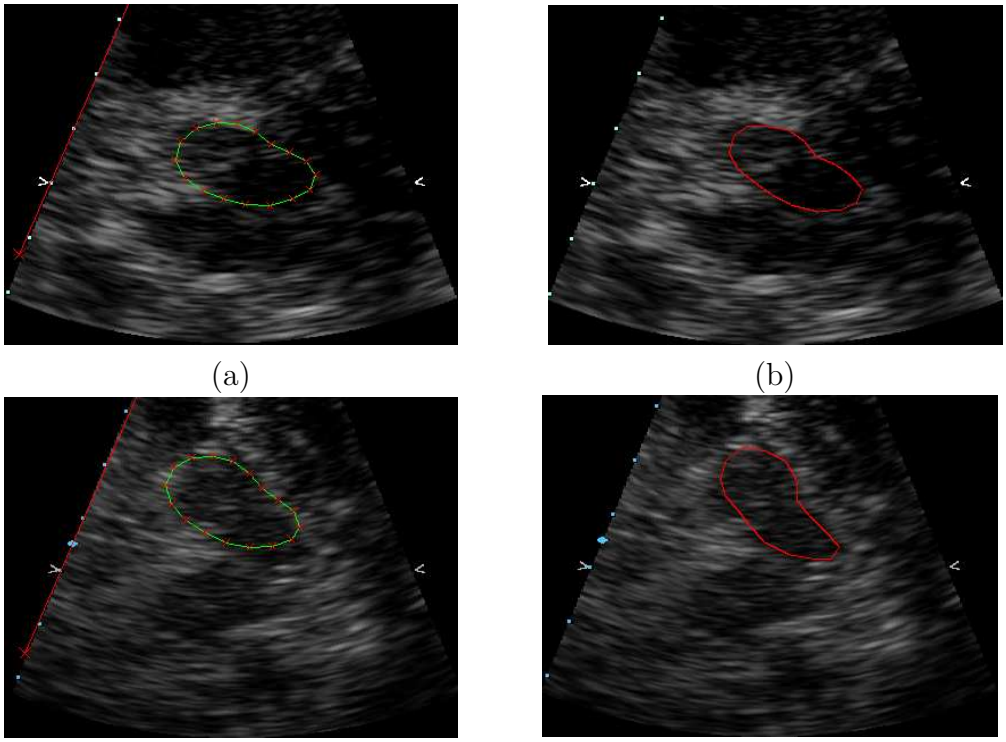
$$I_{new} = \frac{I - I_{min}}{I_{max} - I_{min}} \cdot 255, \tag{2.4}$$

where I_{min} and I_{max} are the minimum and maximum gray values in the patch image, respectively. The TCS image and the histogram are shown in Figure 2.6 (a).

A considerable difference between the original and the normalized images is illustrated in Figure 5.1 in Chapter 5. The second method is zero mean and unit variance normalization, each image is normalized by

$$I_{new} = \frac{I - \mu}{\sigma}, \tag{2.5}$$

where μ and σ are the mean and the standard deviation of the patch image, respectively. As the third alternative, the shape of the image histogram is changed by the contrast-limited adaptive histogram equalization (CLAHE) [55].



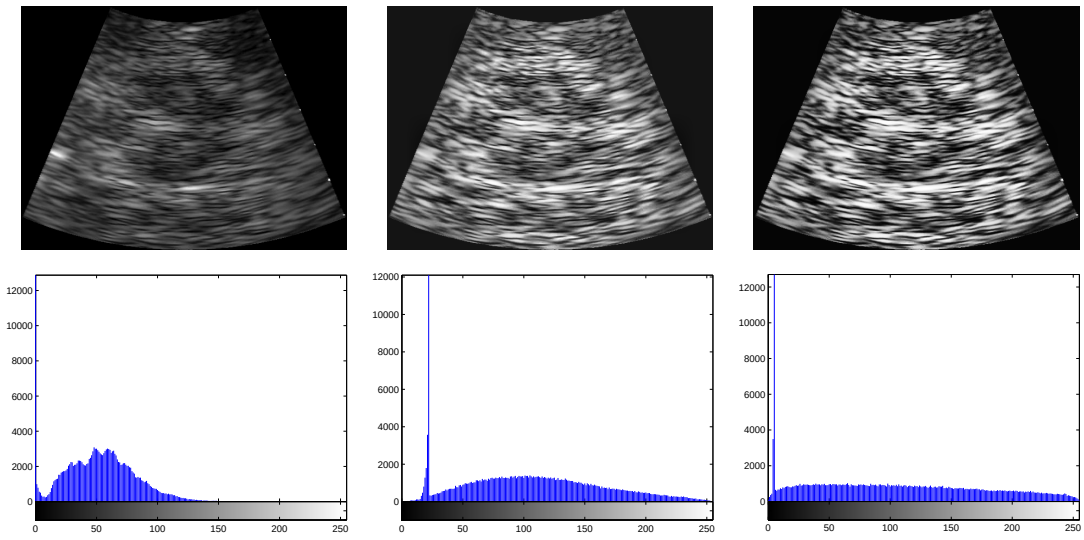
(a) (b) (c) (d)

Figure 2.5: The illustration of the semi-automatic segmentation of midbrain. The upper row and the bottom row are the images of a PD patient and a healthy control subject, respectively. (a) and (c): The initial contours are put on the region of the ipsilateral mesencephalon wing manually; (b) and (d): The segmented results of the proposed segmentation method.

As a results, the histogram of each image is transformed to match with a desired shape. The Rayleigh and exponential distributions are used as the target shapes in this work. The example TCS images are processed by a MATLAB function 'adapthisteq' in Image Processing Toolbox, and the results are shown in Figures 2.6 (b) and (c).

2.3.2 Normalization of the Segmented Region

Regarding the variation of the orientation and shape of the midbrain from one subject to another, one solution is using the rotation-invariant Gabor filter which was investigated in [39], for which the performance was better than the conven-



(a) (b) (c)

Figure 2.6: The original TCS image and the simulated image to the user setting. The upper row are TCS images and the corresponding histogram are in the bottom row. (a) The original TCS image. (b) The shape of histogram is transformed to the bell-shape, Rayleigh distribution. (c) The curved histogram transformed by exponential distribution.

tional Gabor filter for the TCS image classification. Another solution is to align the direction of half of mesencephalon before applying a texture analysis method as mentioned in [32]. Learning from the prior knowledge of the anatomic location of half of mesencephalon and SN [27], a mask is created from the ellipse then fitted onto the ROI. This mask is used for the exclusion of the detected blobs that are outside of the half of mesencephalon region in one previous work [32]. Here, in order to align the half of mesencephalon region of each subject for the same orientation, each manually segmented boundary of half of mesencephalon is fitted with an ellipse and then transformed onto a target ellipse with the affine transformation. The centers of the ellipses and the eight control points on the ellipses are used to calculate the transition matrix of the affine transform. The target ellipse and one fitted ellipse with a ROI are shown in Figures 2.7 (a) and (b), respectively.

Using the transition matrix, the half of mesencephalon and SN areas of the manual segmentations are transformed onto the target ellipse (Figures 2.8 (a) and (c)). The original image of the half of mesencephalon region and its affine adaptation result can be seen in Figures 2.8 (b) and (d), respectively.

(a) (b)

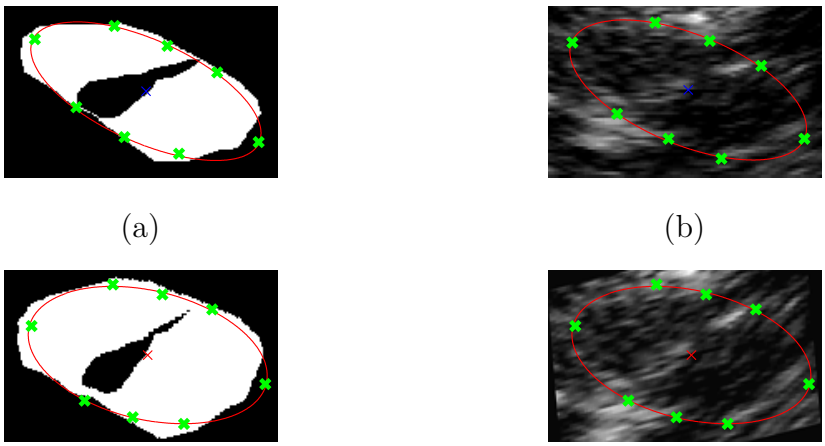
Figure 2.7: The target ellipse and the fitted ellipse. (a) Two green lines are parallel to the minor ellipse axis and across the two ellipse focuses in the target ellipse, respectively. (b) The fitted ellipse with the ROI. Eight green points on the ellipse are used to transform the fitted ellipse into the target ellipse with the affine normalization.

2.4 Experimental Results and Discussion

2.4.1 The Evaluation of The Applied Segmentation Method

The common strategy of evaluating the segmentation methods is using qualitative and quantitative analysis based on the comparison with a gold standard such as the manual segmentation results by experts. Engel et al. evaluated their model-based midbrain and SN segmentation with ten data sets of TCS images [40] and also 30 data sets from cerebral MRI images [48]. The mean of the Hausdorff distance, mean squared distance and region overlap were calculated for comparison with manually segmented data sets by an expert [40]. Sakalauskas et al. applied the same strategy to evaluate their automated midbrain and SN segmentation approach by comparing the segmented contours with the manually marked contours by two experts in 40 images. The calculation of the Hausdorff distance between the contour obtained by their proposed method and the contour outlined by experts was given in [42]. The quantitative results for both midbrain and SN regions were computed for the segmentation method reliably. For 3D data, Ahmadi et al. evaluated their midbrain segmentation algorithm by computing the mesh surface distance map between the output of the proposed method and a ground truth which contains the manually segmented regions, midbrain, SN left and SN right by expert in 11 diagnosed PD patients and 11 healthy controls.

Our aim is to provide a reliable segmentation method for the ROI extraction. The ROIs will be used for feature extraction methods in the next step. Here, we compare the classification results that are based on the feature vectors extracted



(a) (b)

(c) (d)

Figure 2.8: Affine normalization of half of mesencephalon region using a target ellipse. The ellipses and eight control points are superimposed on the segmentations and the original images. (a) manually segmented half of mesencephalon with the fitted ellipse and SN area. (b) The original TCS image of half of mesencephalon with the fitted ellipse. (c) The transformed result with the target/standard ellipse. (d) The affine adaptation result of the original image.

from the labeled data by physicians and the segmented ROIs by the proposed segmentation method. The details about how to calculate these features will be given in Chapter 3 Section 3.1.1. The applied segmentation method was evaluated based on the two datasets that were obtained from Philips SONOS 5500, and the description of these TCS images was mentioned in Section 1.2.3. First, we combined the images of group 1 and group 3 as PD data for the comparison with the healthy controls (group 4). Second, the TCS images in groups 1 and 2 were combined as Parkin mutation data to separate Parkin mutation carriers from the healthy controls. In addition, TCS images obtained with Siemens Sonoline Antares were used to demonstrate the segmentation results as Figure 2.10.

Basically, the idea is to characterize the 'content' of an image histogram using some descriptors. Therefore, the following statistical features [56] of the histogram were calculated for quantitative analysis of the gray-level distribution in the ROI as listed in Table 2.1. The features $F(1, 2, 12, 13, 14)$ were calculated from the intensity values of the original images. To minimize the effect the brightness and the contrast variation due to different user settings, we nor-

malized the TCS images by scaling the gray-level images to a certain range $[0, 64]$. Considering the different values of the window and level adjustment on the ultrasound machine, the ground pixels (gray value 0) were excluded from the calculation of the other features. The motivation and the details will be given in Chapter 3, Section 3.2.

Table 2.1: The statistical features used for the evaluation of the segmentation method.

Feature vector	Feature name
F(1,2)	Mean, variance of ROI
F(3,. . .,7)	the 3rd $\sim$ 7 th order moment
F(8,9)	Normalized value of mean and variance
F(10,11)	Skewness, kurtosis of ROI histogram
F(12)	Root mean square (RMS) contrast [57]
F(13,14)	Skewness, Kurtosis of ROI
F(15,16)	Energy, Entropy of ROI
F(17)	Gray mode, the global max in a histogram

Three feature vectors, PD (66×17), Parkin mutation (58×17) and the healthy control (46×17), are computed from the manually labeled and segmented ROIs, respectively. The performance of these features is evaluated by the sequential feature selection method SFFS. The forward floating search strategy is used to establish the best feature subset by optimizing the criterion function. For SFFS, the criterion function is set as the support vector machines (SVMs) with leave-one-out cross validation method. Regarding the parameters of SVM, the sequential minimal optimization method (SMO) was specified to find the separating hyperplane and the linear kernel function was selected in order to easier analysis of the relationship between the selected features rather than the Gaussian radial basis functions (RBF). The performance of each feature for the two classification tasks are illustrated in Figure 2.9. The classification rates of each feature based on the labeled and segmented ROIs are not exactly similar. We then use the labeled data as the ground truth to evaluate the performance of the applied segmentation approach.

The features extracted from the labeled data were used for the two classification tasks with the four optimal feature subsets (dimension from one to four) that were obtained by SFFS as listed in the second and fourth column in Table 2.2. The same procedure was implemented on the segmented data, and the classification results for four optimal feature subsets are shown in the third and fifth column in Table 2.2. The classification rates for the selected features that were

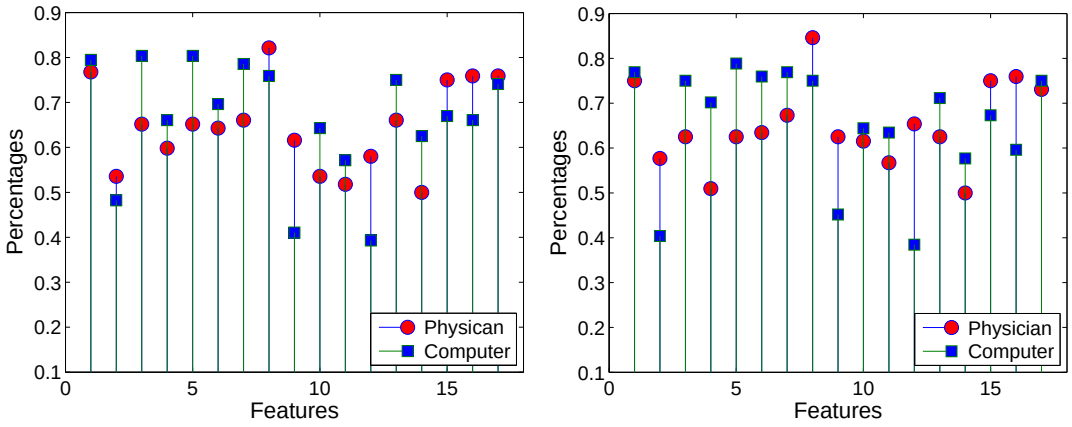


Figure 2.9: The percentages of the classification rates for each statistical feature. Datasets: (a) PD and healthy control; (b) Parkin Mutation and healthy control. The red circles indicate the features calculated from the images labeled by the physicians. The blue squares indicate the classifications based on the applied segmentation method.

calculated from the segmented data are slightly lower than the accuracies of the features computed from the labeled data on each dimension of feature subset. The differences between the labeled and segmented ROIs resulted in a quite similar classification performance based on the feature analysis algorithm. Only one disadvantage of the semi-automatic segmentation is that we need to locate the initial contour manually.

Table 2.2: Feature selection results and the corresponding classification rates (%). 'Physician' indicates the images labeled by physicians and 'Computer' describes the segmentations.

Size of	PD and control data		Mutation and control data	
Feature subset	Physician	Computer	Physician	Computer
1	82.14	80.36	84.62	78.85
2	84.82	83.04	85.58	83.56
3	85.71	83.93	86.54	83.65
4	87.50	83.04	86.54	84.62

After successfully applying the semi-automatic segmentation on the TCS images from Philips SONOS 5500, we then use this segmentation tool to segment the TCS images obtained with Siemens Sonoline Antares. The illustration of the

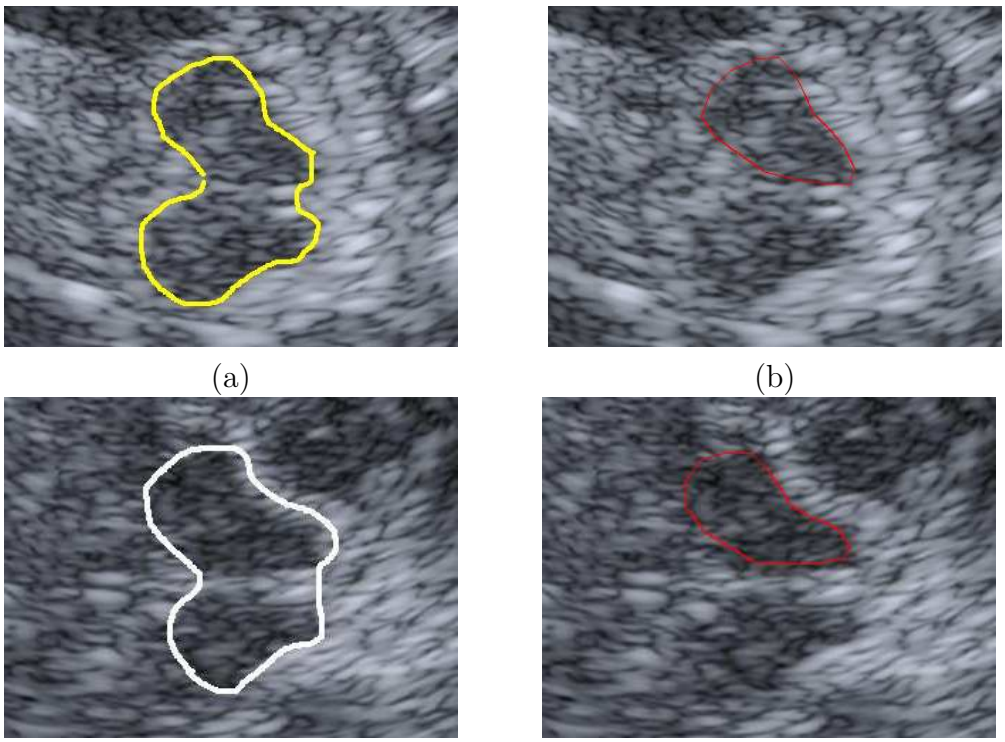


Figure 2.10: The illustration of the semi-automatic segmentation of midbrain on TCS images from Sonoline Antares. The first row and the second row are the images of a PD patient and a healthy control subject, respectively. (a) and (c): The doctor markers on the region of the mesencephalon wings; (b) and (d): The segmented results of the applied segmentation method.

segmentation results are given in Figure 2.10.

2.5 Conclusions

This chapter has given an overview of the segmentation approaches that were applied on TCS image. Actually, due to the property of ultrasound image, it is very difficult to implement a segmentation method that can yield a very accurate output as based on other medical image, such as MRI. Our solution is more suitable for the reduction of the investigator-independence problem than others. We applied a semi-automatic segmentation method to extract the half of mesencephalon, and based on the extracted ROIs, the multiple features were

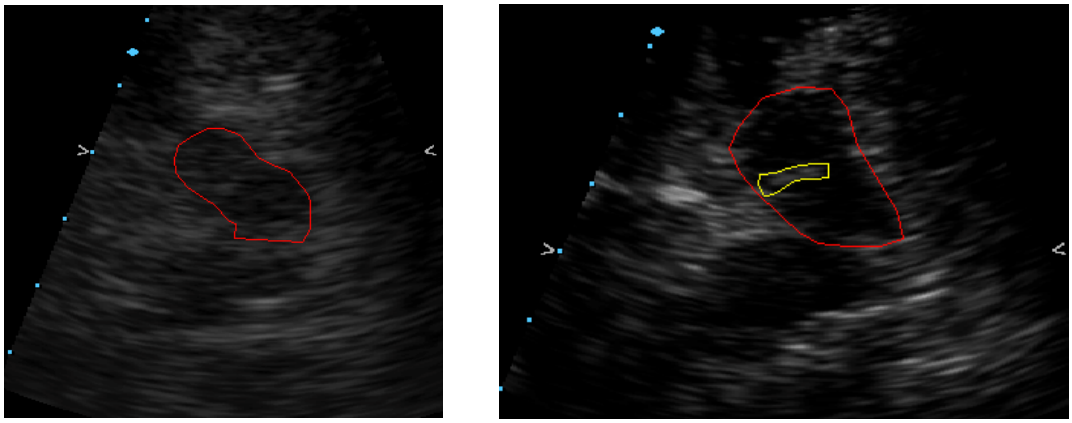
computed for classification. Then, these features were evaluated by feature selection methods. The best feature subset can be used for the classification of TCS images afterwards. Therefore, we avoid the difficulties of accurate segmentation, such as the size of SN hyperechogenicity, meanwhile, we can achieve very good performance of classification by selected features. The details about how to extract multiple features and the feature selection methods will be introduced in next chapter.

Chapter 3

Multiple Feature Extraction From TCS Image

Early diagnosis of Parkinson's disease is of great importance, since clinical symptoms do not occur until the substantia nigra (SN) neurons in the brain stem have been irreparably damaged [27]. Early diagnosis of PD may have two different meanings: the earliest possible PD diagnosis when first motor symptoms are present or preclinical diagnosis of predisposed individuals before first parkinsonian motor symptoms appear [11]. Nowadays, it is possible to determine the formation of idiopathic PD as well as monogenic forms of parkinsonism at an early state by means of TCS [9].

However, this finding is still subject to manual evaluation of the examined images. For quantitative analysis of SN hyperechogenicity, only the area of SN rather than the other image characteristics have been considered. Our goal is to reduce investigator-dependence of the diagnosis by extracting multiple features from the manually segmented ipsilateral mesencephalon wing, which is close to the Ultrasound probe as shown in Figure 3.1. The moment of inertia and Hu1-moment were found by Kier et al. [27] as good parameters for separating control subjects from parkin mutation carriers. In this chapter, we propose hybrid feature extraction methods which include statistical, geometrical and texture features for the early PD risk assessment. These features are used with an SVM classifier. The performance of SVMs often does not increase with the growth of the feature set, therefore the feature selection methods such as sequential backward selection (SBS) and sequential forward selection (SFS) are applied to obtain the best feature subset.



(a) (b)

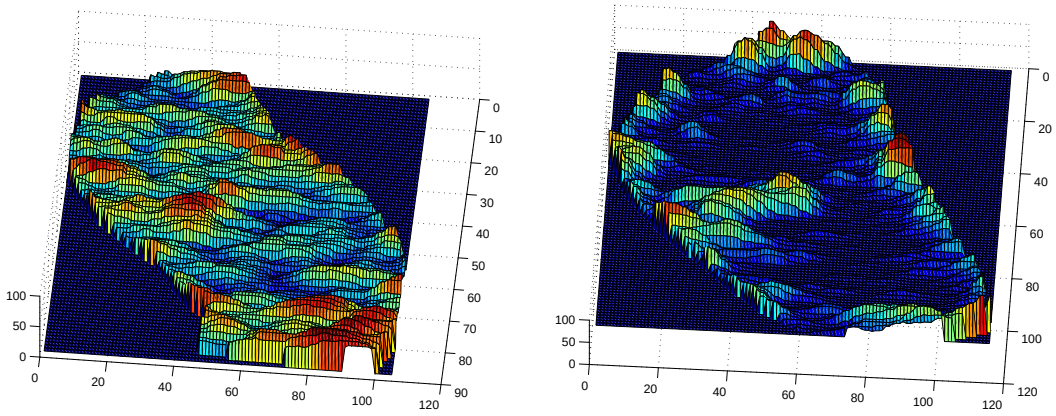
Figure 3.1: Manually segmented upper half of the mesencephalon (red) in (a) a healthy subject and (b) a PD affected subject, SN region is indicated by a yellow contour.

3.1 Multiple Feature Extraction

The feature extraction is used to reduce the dimension of the input data and minimize the training time taken by the classifier. Multiple features which include geometrical moments, statistical moments and texture moments are extracted from the region of interest (ROI). The ultrasound images of the upper half of the mesencephalon, ROIs, are shown in Figure 3.3 (a) and (c), which are manually segmented from Figure 3.1 by physician. The difference of gray value distributions between these two class images can be seen more clearly in the surface plot (Figure 3.2) than in a gray image.

3.1.1 Statistical Features

Feature-specific measurements in images can be divided into four general classes: size, shape, brightness, and location [58]. Within each class, a variety of specific measurements and operations can be performed. Normally a numeric output produced by measurements is suitable for statistical analysis. The first-order statistics measure the likelihood of a gray value at a random location in image. It can be calculated from the histogram of pixel intensities [59]. The basic idea is to characterize the 'content' of an image histogram using some descriptors. The two histogram-related features, energy E and entropy H can be defined. The energy



(a) (b)

Figure 3.2: Surface plot of the ROI in (a) Healthy subject and (b) PD affected subject. (two same subjects with Figure 3.1)

assumes its maximum value close to one if the image has a histogram with only one bin. The value of the energy deceases when the image has a broader intensity variations. The entropy is a measure of information content which is inversely related to the energy in a qualitative manner. For example, a predominantly random distribution has a high entropy and uniform distribution has a low entropy. Moreover, two shape-descriptors of a probability distribution, skewness and kurtosis are used to characterize the shape of the histogram. Skewness is a measure of the extent of the asymmetry to probability distribution around the mean. Kurtosis is a measure of how outlier-prone a distribution is, compared to the normal distribution (kurtosis value is 3). In this section, the histograms that are more outlier-prone than the normal distribution have kurtosis values higher than 3. In addition, the gray mode is a measure of the peak, the global maximum of a histogram.

Therefore, the following statistical features [56] of the histogram are calculated for quantitative analysis of the gray-level distribution in the ROI as listed in Table 3.1.

3.1.2 Geometrical Features

In the fields of pattern recognition and computer vision, a variety of approaches have been implemented for recognizing and matching objects by their attributes and geometrical feature [60]. Instead of shape descriptors perimeter, such as area

Table 3.1: 15 statistical features are extracted from ROIs.

Feature vector	Feature name
F(1,2)	Mean, variance of ROI
F(3)	the 3rd order moment
⋮	⋮
F(10)	the 10 th order moment
F(11,12)	Energy, Entropy of ROI
F(13,14)	Skewness, Kurtosis of ROI
F(15)	Gray mode, the global max in a histogram

and compactness, moment analysis represents a systematic method. Moment analysis based on region-interior pixels, such as the evaluation of central moments, normalized central moments, and moment invariants. The quantities are suitable for object recognition due to the shape attributes that are independent of object size and orientation [60].

Given a binary image, region-interior pixels are the value 1, and the moments m_{pq} of a region can be defined as:

$$m_{pq} = \sum_{R} x^p y^q, \tag{3.1}$$

where the sum runs over all region-interior pixels. The expressions of the moments m_{pq} can be found in Chapter 4 in [60]. The moments m_{pq} simply represent measurements of area, but the area cannot be scale invariant because the objects may depend on the scale of the image. One solution is the normalization by the area, the central moments μ^{pq} are defined in terms of the centroid location:

$$\mu_{pq} = \sum_{R} (x - x_c)^p (y - y_c)^q. \tag{3.2}$$

The central moments are usually divided by the zeroth moment to generate the normalized central moments. A list of the central moments and moment invariants is given in the chapter 4 in [60].

Christian Kier et al. found both the moment of inertia and Hu1-moment could be used to separate control subjects from parkin mutation carriers [27]. Seven moments defined by Hu [61] were computed based on the segmented ROIs. Hu moments have been proven to be invariant to object scale, position and orientation. The moment of inertia is adapted to image processing by interpreting intensity values as inertia values and varies strongly between a uniform and a centrical distribution of the ROI [27]. The geometrical moments used here are listed in Table 3.2:

Table 3.2: A total of 8 moment features are extracted from ROIs.

Feature vector	Feature name
F(16)	Moment of inertia
F(17)	the 1st order Hu moment
⋮	⋮
F(23)	the 7th order Hu moment

3.1.3 Texture Features

Looking at a picture of real objects, we can use our vision system to connect regions having a certain similarity, such as similar gray or color value, into objects. But, these images of objects often do not show uniform intensities in regions. For example, a grass image or a brick wall image. We can recognize that these regions are not uniform but contain variations of intensities, these certain repeated patterns can be called visual texture.

While it is hard to give a formal definition of texture, such as the repetition of a pattern or patterns over a region [62], a region in an image has a constant texture if a set of local statistics or other local properties are constant or approximately periodic [63]. Texture analysis is an important area of study in practical machine vision applications [59]. Texture analysis methods have been implemented in a variety of applications, such as automated inspection and medical image analysis. The analysis techniques include extraction procedure of features from the medical image, the classification tasks based on the extracted features. Depending on the specific application, the extracted features capture morphological properties and textural properties of the image [59].

An often used wavelet in image analysis is the Gabor wavelet, $G(x, y)$. In 2D, it is the product of a Gaussian with a plane wave [60]. Gabor wavelet analysis produces images over many orientations for a single spatial frequency, it localizes features in the spatial domain compared with frequency domain analysis. In this chapter, we propose to use Gabor filter for the texture analysis of TCS image. The illustration of the responses of Gabor filter on TCS images are shown in Figure 3.3. Another texture analysis method characterizes the region properties in its gray levels by computing the co-occurrence matrix. Grey-Level Co-occurrence Matrix (GLCM) texture measurements were proposed by Haralick [64], who formulated 14 texture metrics based on the co-occurrence matrix [65]. The details about GLCM will be given in Section 3.1.3.

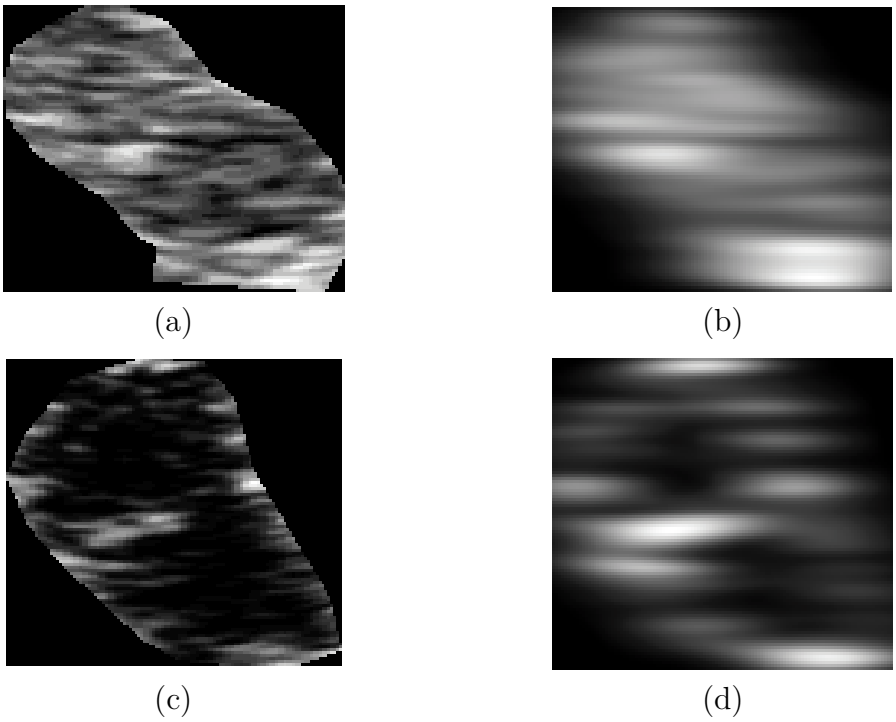


Figure 3.3: The two TCS images are selected from a healthy and a PD affected subject and the corresponding filtered image by Gabor filter (scale 0, orientation 1). (a) Half of the mesencephalon (ROI) in TCS image of a healthy subject and (c) PD affected patient. (b) and (d) are the corresponding Gabor filter processing results.

Gabor filter

In this section, texture features that are extracted by a bank of Gabor filters from the region of interest (ROI) are shown in Figures 3.3 (a) and (c), which were manually segmented from Figure 3.1 by physicians. The Gabor filter (scale 0, orientation 1) processing results are given in Figures 3.3 (b) and (d).

Given an image $I(x, y)$ with size $P \times Q$, its discrete Gabor wavelet transform is then defined by a convolution:

$$G_{mn}(x, y) = \sum_{\xi} \sum_{\eta} I(x - \xi, y - \eta) \mathbf{g}^*_{mn}(\xi, \eta), \tag{3.3}$$

where $*$ indicates the complex conjugate of g_{mn} [66]. The filter mask size is

indicated by ξ and η. The two dimensional Gabor function $g(\xi, \eta)$ is:

$$g(\xi,\eta) = \frac{1}{2\pi\sigma_\xi\sigma_\eta}\exp[-\frac{1}{2}(\frac{\xi^2}{\sigma_\xi^2}+\frac{\eta^2}{\sigma_\eta^2})]\cdot\exp[2\pi jW\xi], \tag{3.4}$$

where W is called the modulation frequency, and ξ and η range from -30 to 30, the filter mask size is 61×61. The generating function is

$$g(\xi,\eta) = a^{-m}\mathbf{g}_{mn}(\tilde{\xi},\tilde{\eta}), \tag{3.5}$$

and

$$\tilde{\xi} = a^{-m}(\xi cos\theta + \eta sin\theta); \qquad \tilde{\eta} = a^{-m}(-\xi sin\theta + \eta cos\theta), \tag{3.6}$$

where m and n specify the scale and orientation respectively, $a > 1$ and $\theta = n\pi/N$. N is the total number of orientations. Moreover,

$$a = (U_h/U_l)^{\frac{1}{M-1}}; \qquad W_{m,n} = a^m U_l; \tag{3.7}$$

$$\sigma_\xi = \frac{(a+1)\sqrt{2\ln 2}}{2\pi a^m(a-1)U_l}; \qquad \sigma_\eta = \frac{1}{2\pi tan(\frac{\pi}{2N})\sqrt{\frac{U_h^2}{2\ln 2}-(\frac{1}{2\pi\sigma_\xi})^2}}. \tag{3.8}$$

It is assumed that the SN region in the ROI (half of mesencephalon) has homogeneous texture, therefore the mean μ_{mn} and the standard deviation σ_{mn} of the transform coefficients' magnitudes are used to represent the texture features:

$$\mu_{mn} = \frac{\sum_x\sum_y \mid G_{mn}(x,y) \mid}{P\times Q}, \tag{3.9}$$

$$\sigma_{mn} = \frac{\sqrt{\sum_x\sum_y(\mid G_{mn}(x,y) \mid -\mu_{mn})^2}}{P\times Q}. \tag{3.10}$$

The Gabor feature vector f is composed by μ_{mn} and σ_{mn} as feature components [67]. Five scales and six orientations have been used in the experiments:

$$f = (\mu_{00},\sigma_{00},\mu_{01},\sigma_{01},...,\mu_{45},\sigma_{45}). \tag{3.11}$$

Based on the designed Gabor filter bank, the extracted Gabor features are listed in Table 3.3. The other two texture features, average gray level and average contrast, are computed as in [68].

Table 3.3: The Gabor filter bank designed with five scales and six orientations. Two conventional Gabor features, mean and standard deviation are extracted based on the magnitude of each filter response. A total of 60 Gabor features.

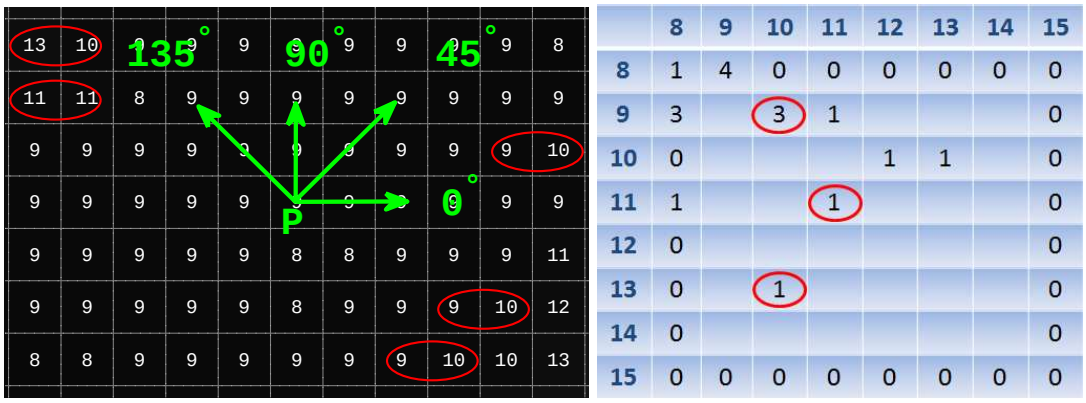
Feature vector	Feature name
F(24)	Average gray level
F(25)	Average contrast
F(26)	Mean of the magnitude $f(1)$
F(27)	Standard deviation of the magnitude $f(2)$
⋮	⋮
F(84)	Mean of the magnitude $f(59)$
F(85)	Standard deviation of the magnitude $f(60)$

GLCM features

A region has a particular texture in an image, the properties of the variation in its gray levels can be used to identify the region [69]. Then a texture measurement is required to reflect the property about the repeating nature of the texture in the image, such as repeating horizontal and/or vertical lines with similar widths and separations. A gray-level co-occurrence matrix is suitable to provide the information about the positions of pixels with similar gray-level values. The idea is to scan the image and record the information about how often a pixel pair in value appears that differ by a fixed distance in position. A definition of GLCM $C_d^\theta(i,j)$ is given as how often different combinations of a pixel pair (i,j) occur in an image, that located at a distance d and direction θ. The term $C_d^\theta(i,j)$ is a two-dimensional histogram of gray values i and j, and the element at (i,j) is the probability of this pixel pair. Furthermore, Gray level co-occurrence probability is defined as the number of times this combination occurs, divided by the total number of possible outcomes. The normalization equation reads:

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i,j}^{N-1} V_{i,j}}, \tag{3.12}$$

where i,j are row and column number, respectively, $V_{i,j}$ is the combination occurs. For example, for an image with 256 gray levels, the size of the GLCM matrix $C_d^\theta(i,j)$ is 256×256. In the experiment, the multiple GLCMs were created with four directions ($\theta = (0°, 45°, 90°, 135°)$), the window size was chosen as 3 by 3 (offset size is one). An illustration regarding a patch image and the pixel of interest with four direction is shown in Figure 3.4 (a). The correspond-



(a) (b)

Figure 3.4: The illustration of creating a Gray-Level Co-Occurrence Matrix. (a) A patch image describes the spatial relationships of pixels that are defined by offset (the distance from the pixel of interest P, set as 3) and four directions. (b) A part of the GLCM matrix, the elements contain the appeared times of the instances, such as $(13, 10)$, $(11, 11)$, and $(9, 10)$ in the patch image (a).

ing GLCM matrix with $\theta = 0°$ and some elements are shown in Figure 3.4 (b). The values of the combinations $V_{13,10}$, $V_{11,11}$, and $V_{9,10}$ are 1, 1, and 3, respectively, which are the numbers of times these three instances appear in the patch image.

Haralick et al. have defined 14 global metrics to classify the texture, and only four key measures are introduced in this section. The GLCM feature vector g is composed by four features, such as contrast (inertia), correlation, energy (angular second moment), and homogeneity. The contrast and homogeneity are approximately the inverse of each other.

The GLCM feature, energy, measures the uniformity in the GLCM matrix. The definition is given as

$$gl_{u1} = \sum_{i,j=0}^{N-1} P_{i,j}^2. \tag{3.13}$$

The GLCM feature, contrast, is a measurement of the local variations in the GLCM matrix:

$$gl_{u2} = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2. \tag{3.14}$$

Correlation measures linear dependency of gray levels pairs. In another words, it measures the joint probability occurrence of the specified pixel pairs:

$$gl_{u3} = \sum_{i,j=0}^{N-1} P_{i,j} \frac{(i-\mu_i)(j-\mu_j)}{\sqrt{\sigma_i^2 \sigma_j^2}}. \tag{3.15}$$

Homogeneity, inverse difference moments, measures the inverses of the contrast. In another words, it measures the closeness of the distribution of combination $V_{i,j}$ to the GLCM diagonal:

$$gl_{u4} = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}. \tag{3.16}$$

After creating the GLCMs, the four statistical features are computed which provide information about the texture of an image. The GLCM feature vector g consists of 16 GLCM features gl_{uv} with four directions u and four features v at each direction.

$$g = (gl_{11}, gl_{12}, ..., gl_{43}, gl_{44}), \tag{3.17}$$

Until now, the feature vector F created by multiple feature extraction method has 101 dimensions, the features from GLCM are listed in Table 3.4.

Table 3.4: Final feature vector has 101 features, which includes 16 GLCM features with four directions.

Feature vector	Feature name
F(86)	Energy with direction 1, $g(1)$
F(87)	Contrast with direction 1, $g(2)$
F(88)	Correlation with direction 1, $g(3)$
F(89)	Homogeneity with direction 1, $g(4)$
⋮	⋮
F(98)	Energy with direction 4, $g(12)$
F(99)	Contrast with direction 4, $g(13)$
F(100)	Correlation with direction 4, $g(15)$
F(101)	Homogeneity with direction 4, $g(16)$

3.2 Robust Feature Extraction

In feature extraction, the invariance property means the extraction process does not vary based on the specified conditions [70]. In this section, the robust feature extraction should yield the features reliably and robustly whatever the value of the parameters of the input images. The properties of the TCS images, such as the contrast and brightness, are affected by different settings of the US machine used by different examiners. Furthermore, the challenge of the TCS images classification by using Gabor filters is that the orientations and shapes of the half of mesencephalon are different from one PD patient to another, as shown in Figure 3.5.

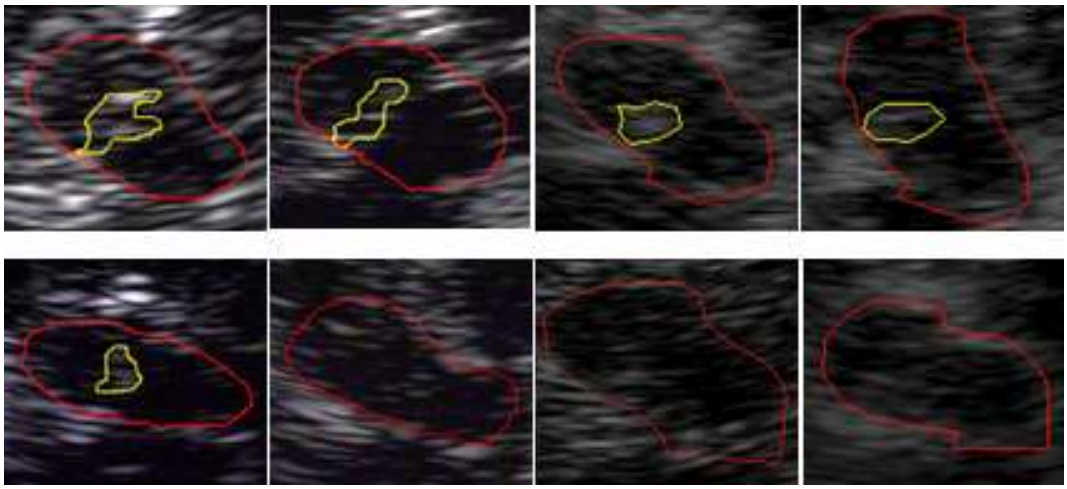


Figure 3.5: Manually segmented TCS images by physicians from PD patients and the healthy control as illustrated in the upper and bottom row, respectively. All images are collected from Philips SONOS 5500.

3.2.1 Features of the Normalized Histogram

Observing the TCS images, the range of the gray levels is different from one image to another due to the different user settings, such as the parameters of the window level and window center. Therefore, the ROIs need to be normalized before the feature extraction procedure. The histogram of each ROI is normalized or equalized from the range of $[0, 255]$ to $[0, 64]$, then the histograms that belong to the same group are added up and the mean of each summation is illustrated in Figure 3.6. The histograms of healthy controls in Dataset 1 (group 4) are obviously different from others. Moreover, the background (pixel values close to

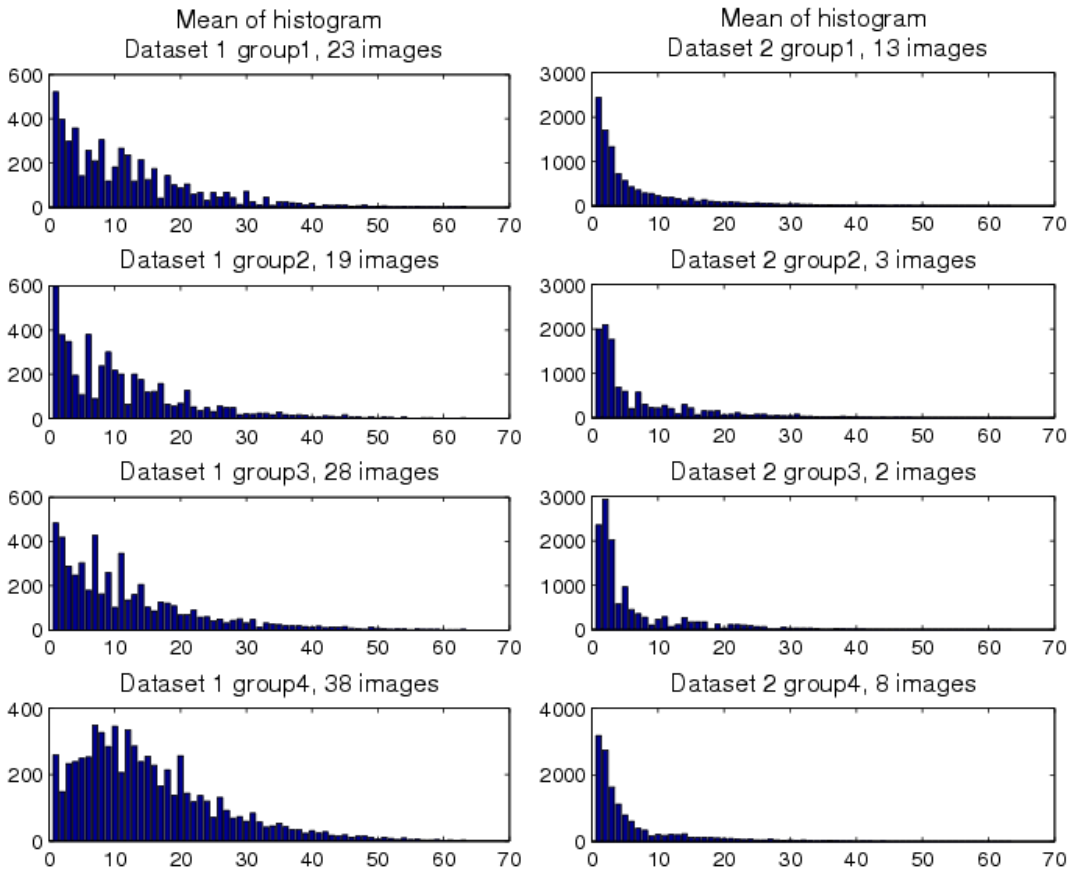


Figure 3.6: The illustration for the averaged histograms that are calculated from ROIs in (a) Dataset 1 and (b) Dataset 2, respectively. The histograms for every group are individually shown in each row.

zero) in a gray level image is not stable when a contrast adjustment is applied to the image. Thus, the bin of zero in the normalized histogram is ignored during the feature calculation. As a result, these features of the normalized histogram are less sensitive to the contrast changes than the statistical features in Section 3.1.1.

3.2.2 Rotation-invariant Gabor Features

As shown in Figure 3.5, the orientation and the shape of each ROI (half of mesencephalon) are different. In this section, our goal is to develop Gabor features that are invariant to the direction of ROI, the brightness, and the

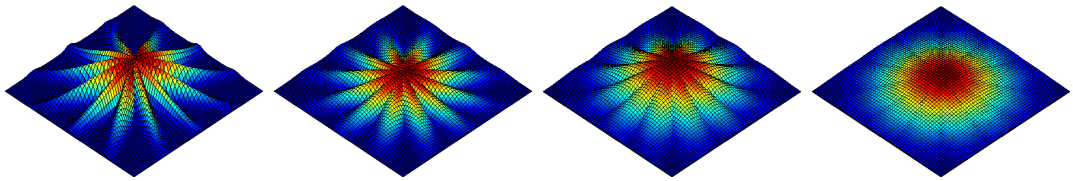
contrast changes. Therefore, we propose a texture analysis method that applies a rotation-invariant Gabor filter bank on the half of mesencephalon area and computes the histogram features from the filtered images for the TCS image classification. Two methods are implemented for solving the rotation-invariant problem. The first one is to combine all the Gabor filter response linearly with different orientations, but at each scale level [71]. Another method was proposed in [72] which shifting circularly the filtered image at same scale to achieve the rotation-invariance property.

Rotation-invariant Gabor Filter Bank Design

Han et al. claimed that the summation of the filter responses with different orientations could yield a rotation-invariant Gabor filter [71]:

$$g_m^{(R)}(\xi,\eta) = \sum_{n=0}^{K-1} g_{m,n}(\xi,\eta), m = 0, 1, ..., S-1. \tag{3.18}$$

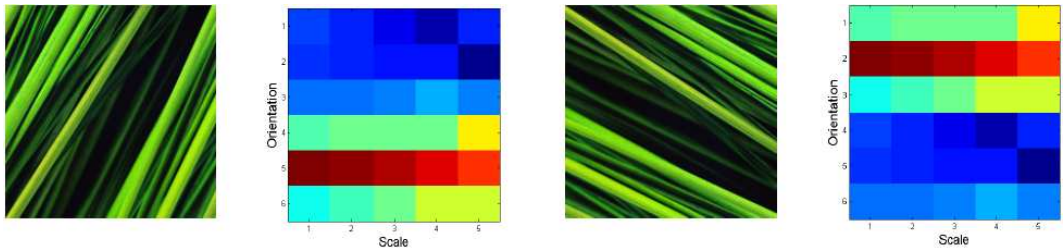
The illustrations of the summation method with four scales and six orientations are shown in Figure 3.7.



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

Figure 3.7: Summing all the Gabor filter response with different orientations at the same scale level. From left to right, the Gabor filter responses are summed at scale 1, 2, 3, and 4, respectively.

Second, The Gabor features can be sorted by the total energy of the corresponding filtered images over the orientation at the same scale. Therefore, the Gabor features shifted by the dominant direction could be independent to 2D rotation. The dominant direction [72] at each scale is defined as the direction with the highest energy. An illustration of the dominant direction in the energy map is given with an example texture image in Figure 3.8. The energy of each filtered image is calculated for the circularly shifting. The filtered image G_{ij} with the dominant direction j is moved to be the first position at scale i, and the other filtered images at scale i are circularly sifted accordingly. While, the



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

Figure 3.8: The energy maps of the texture image and the rotated image. The dominant direction is indicated as red color in energy maps (b) and (d). (a) The example grass image and (b) the corresponding energy map. (c) The rotated grass image (90°) and (d) the corresponding energy map.

feature elements μ_{ij}, σ_{ij} (in Equation 3.9 and Equation 3.10, respectively) of G_{ij} are put at the beginning in the Gabor feature vector f at scale i. At last, a rotation-invariant Gabor feature vector f_r is obtained until all filtered images shifted at their corresponding scales. As a result, the feature elements μ_{ij}, σ_{ij} in conventional Gabor feature vector f are shifted as in the rotation-invariant Gabor feature vector f_r. For example, if f is (A,B,C,D,E,F) and (C) is the dominant direction, then f_r is (C,D,E,F,A,B).

Contrast-invariant Gabor Features

In general, the conventional Gabor features, the mean and the standard deviation are calculated from the intensity values of the filtered image directly. In this section, we compute the entropy from the histogram of the filtered image. The Shannon entropy can be used to measure the randomness from the image histogram [73]. In other words, the entropy measures the uniformity of the filtered image.

The Shannon entropy is given by:

$$HX = \sum_{i=1}^{n} p(x_i) \log_b p(x_i), \tag{3.19}$$

where $Pr[X = x_i] = p(x_i)$ is the probability mass distribution of the signal [74]. Actually, the symbol alphabet of filtered image is in general not finite.

Therefore, a proper measurement of entropy is differential entropy. The entropy of the histogram image is given as:

$$H(X) = -\sum_{x \in S} hist_{norm}(x) \log_2(hist_{norm}(x)), \tag{3.20}$$

where S is the support set of the random variable x and $hist_{norm}(x)$ is the histogram properly normalized to fit a probability density function [73]. The summation of the probability density function $hist_{norm}(x)$ is one.

3.3 Feature Selection

For the classification, the motivation of feature selection is to make a prediction of unknown data by a hypothesis constructed from a certain number of training instances [75]. A hypothesis is a pattern that predicts classes based on learning from given data/instances. The key factor is the number of features that determines the hypothesis space. In other words, feature selection is an essential task to remove redundant features. Redundant features are a type of irrelevant feature [76] that can be removed without influencing learning performance. The selection can be achieved by ranking features according to criterion functions and then selecting the top k features. The other way is to select a minimum subset of features without learning knowledge [75]. The aspects of feature selection include models, feature performance measures, and search strategies. The three typical models are embedded models (an example, the decision tree induction algorithm), filters, and wrappers. In a wrapper model, one uses a learning algorithm, and based on the performance, determines the quality of selected features. The filter models use an approximate measure, such as mutual information, to replace the accuracy to rank the selected features. Compared to the wrapper models, which are computationally intensive, the filter models are faster to compute. In this chapter, the goal of feature selection is to select the best feature subset automatically for TCS image classification purpose by given a feature vector. The SVM classifier has been chosen to evaluate the effectiveness of feature subsets. The feature selection detects an optimal feature subset based on the feature vector F.

3.3.1 Sequential Feature Selection

A general feature selection method, sequential feature selection includes two components. One is a criterion function, which is used to minimize over all

possible feature subsets. In this work, the misclassification rate of SVMs is set as the criterion. Another component is a sequential search strategy, which evaluates the criterion to establish the best feature subset. Search strategies include forward, backward, floating, branch-and-bound, and randomized selection [77]. For the sequential forward selection (SFS), features are selected successively by adding the locally best feature, which is the one that provides the lowest criterion value, to an empty candidate set. The SFS technique starts from the best individual feature (BIF) by identifying the first feature that has the highest discrimination power. The SFS stops until the further features do not decrease the criterion. The procedure of SFS is explained in a flow chart as shown in Figure 3.9.

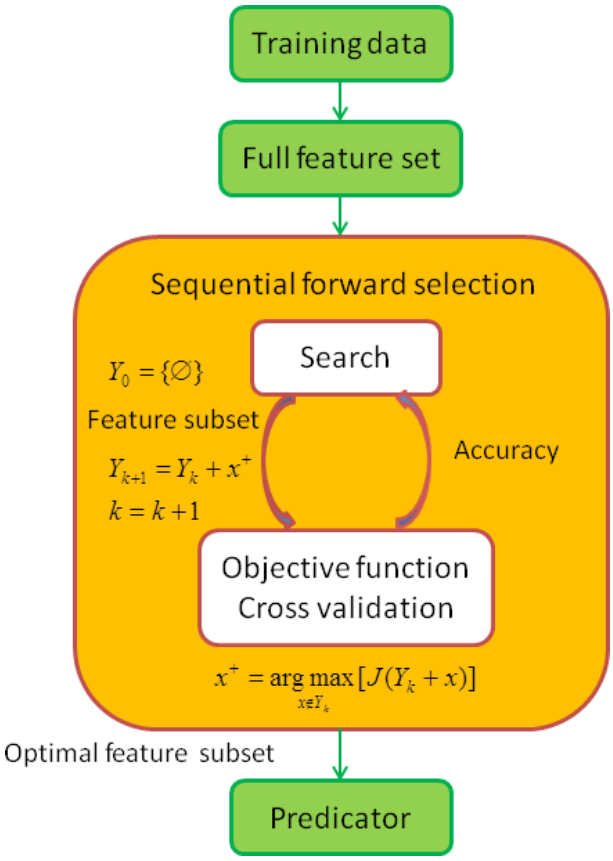


Figure 3.9: The illustration of the sequential forward selection method.

The sequential backward selection (SBS) method is the 'bottom up' counterpart to SFS. In SBS, starting from a full candidate set, one sequentially removes the feature which has the highest criterion until the removal of any further features may lead to an increase of the misclassification rate.

3.3.2 Floating Search Selection

Sequential feature selection is a common feature selection method which includes two search algorithms, SFS and SBS. However these two methods are generally suboptimal and have a disadvantage, the 'nesting effect' [78]. In the case of SFS, the selected features cannot be discarded afterwards. Similar to SFS, the discarded features cannot be re-selected in the case of SBS. As a result, the optimal feature subset might be only suboptimal [78]. Therefore Pudil et al. proposed the floating search methods and found that the performance of the floating search method was better than of others, and the computational complexity was reduced. The procedure of SFFS is explained in a flow chart as shown in Figure 3.10.

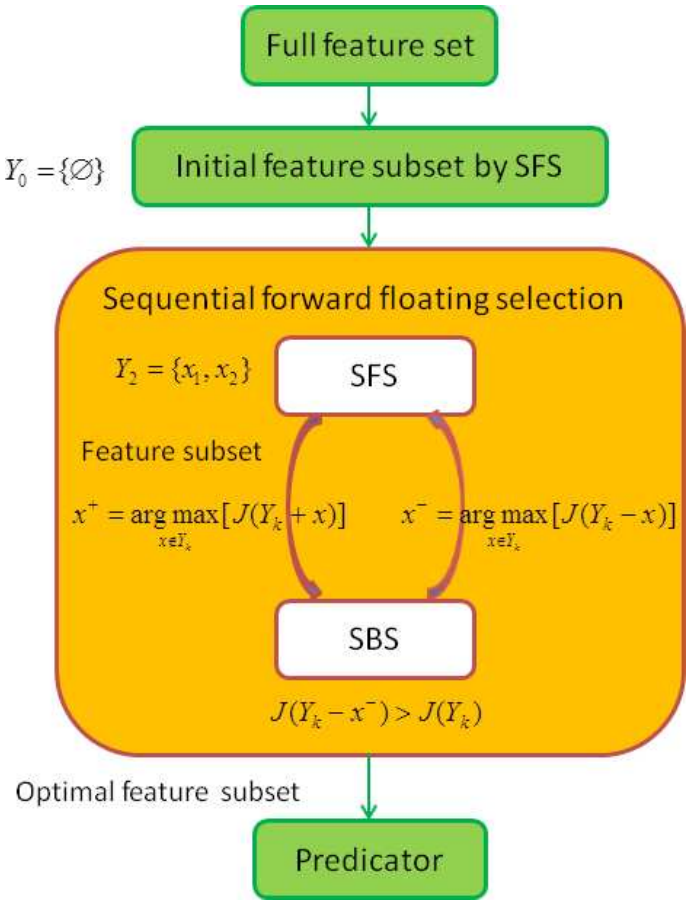


Figure 3.10: The explanation of sequential forward floating selection method.

The character of 'floating' means the selected features can be flexibly changed instead of being fixed. According to the dominant search direction, the floating

search methods consist of sequential forward floating selection in the forward direction and sequential backward floating selection (SBFS) in the opposite direction. Considering the computation time, SFFS characterized by a dynamical changing of features at each step was implemented in this thesis. In this chapter, the misclassification rate of SVMs was set as the criterion, the Gaussian radial basis functions (RBF) was used to map the training data into the kernel space. The sequential minimal optimization method (SMO) was specified to find the separating hyperplane. SFFS was shown to give good results and to be more effective than the SBS and SFS in the Section 3.4.

3.4 Experimental Results and Discussion

We aim at developing the features that can be used to recognize the TCS images (half of mesencephalon) of two different classes called 'healthy controls' and 'Parkin mutation carriers'. The experiment steps are explained in the flowchart in Figure 3.11.

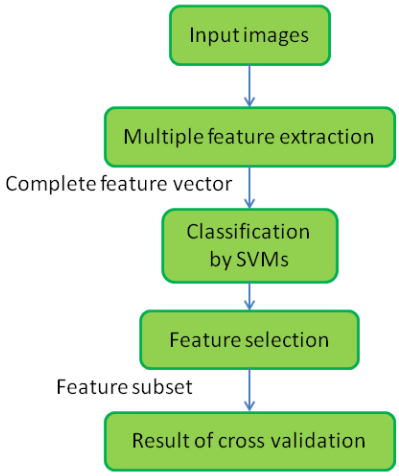


Figure 3.11: The workflow for the TCS images classification based on multiple features.

A clinical study has been conducted to evaluate whether the image features can be used as an early PD indicator. The study consisted of 36 images from 21 healthy controls (subjects without mutation and symptoms of PD) and 42 image from 19 Parkin mutation carriers. All these 40 subjects underwent a detailed neurological examination. Therefore the diagnosis result can be considered as

the ground truth to compare and evaluate the classification. In each image the half of mesencephalon (ROI marked by red contour) and even the SN (yellow contour) area were manually segmented by two individual physicians as shown in Figure 3.1.

3.4.1 Performance of Multiple Features

The multiple features composed of statistical moments, geometrical moments and Gabor texture features that extracted from the ROIs. The SVMs classification has been cross-validated by the leave-one-out method. This gives the accuracies of 79.49%, 79.49% and 76.92% when statistical features, geometrical features and Gabor texture features are added successively into the feature vector. The feauture vector increased while the classification rate did not increase accordingly. Then the SBS and the SFS were used, respectively, to minimize the best feature subset. Comparatively, the feature subset obtained by SFS gives the highest classification rate of 96.15%. In this feature subset, the Gabor features $f(1), f(2)$ have the best individual performance of 88.46%. The detailed results of implementation of these feature sets are given in Table 3.5. Except for the comparison of SFS and SBS, the performance of the individual feature, the first Hu mement, contrast, and two Gabor features are also shown in Table 3.5. Due to the good performance of Gabor features, more texture features (GLCMs) will be computed in next experiment.

Table 3.5: Feature selection results based on multiple features (85 dimensions), classification rates (%) of SVMs with cross-validation. The performance of each feature in the subset seleted by SFS are compared individually [37].

Feature sets	Accuracy	Specificity	Sensitivity
$F(1, ..., 15)$	79.49	66.67	90.48
$F(1, ..., 23)$	79.49	66.67	90.48
$F(1, ..., 85)$	76.92	83.37	71.43
SBS,$F(12, 27, 80, 82, ..., 85)$	92.31	86.11	97.62
SFS,$F(17, 25, 26, 27, 29)$	96.15	94.44	97.62
Hu1,$F(17)$	83.33	80.56	85.71
Contrast,$F(25)$	58.97	61.11	57.14
Gabor,$f(1), f(2)$	88.46	80.56	95.24
Gabor,$f(4)$	87.18	77.78	95.24

3.4.2 Evaluation of Texture Features

The hybrid feature extraction has shown very good performance of TCS classification, especially Gabor features were better than the others in last Section 3.4.3. Therefore, the GLCM texture features are computed and combined with Gabor features. They are then evaluated by feature selection SFFS. The study includes 36 images from 21 controls (subjects without mutation and symptoms of PD) and 42 images from 19 Parkin mutation carriers. All these 40 subjects underwent a detailed neurological examination. In this study, the feature vector F has 85 dimensions, which consists of Hu moments, average gray level, average contrast, Gabor feature vector f and GLCM feature vector g. The features F1 to F85 were extracted as listed in Table 3.6:

Table 3.6: A total of 85 features for the texture analysis, which includes Hu moments, Gabor features, and GLCM features.

Feature vector	Feature name
F(1)	the 1st order Hu moment
⋮	⋮
F(7)	the 7th order Hu moment
F(8)	Average gray level
F(9)	Average contrast
F(10)	Gabor texture feature 1, $f(1)$
⋮	⋮
F(69)	Gabor texture feature 60, $f(60)$
F(70)	GLCM texture feature 1, $g(1)$
F(85)	GLCM texture feature 16, $g(16)$

The performance of texture features were compared with other features, such as Hu moments. The SVMs classification was cross validated by the leave-one-out method. This gave the accuracies of 90.91% and 92.73% when SFS and SFFS were used, respectively, to minimize the best feature subset. We could not obtain a small feature subset by SBS. The feature subset $F(12, 77)$ obtained by SFFS gave the highest classification rate of 92.73% ($F(17)$, Gabor feature $f(3)$ and $F(77)$ is GLCM feature $g(8)$). In this feature subset, the GLCM features $F(73, 77)$ had a good preformance of 90.91%. The detailed results of the implementation of these feature sets are given in Table 3.7. The performance of the optimal feature subset $F(12, 77)$ seperating the TCS image into two classes are illstrated in Figuire 3.12.

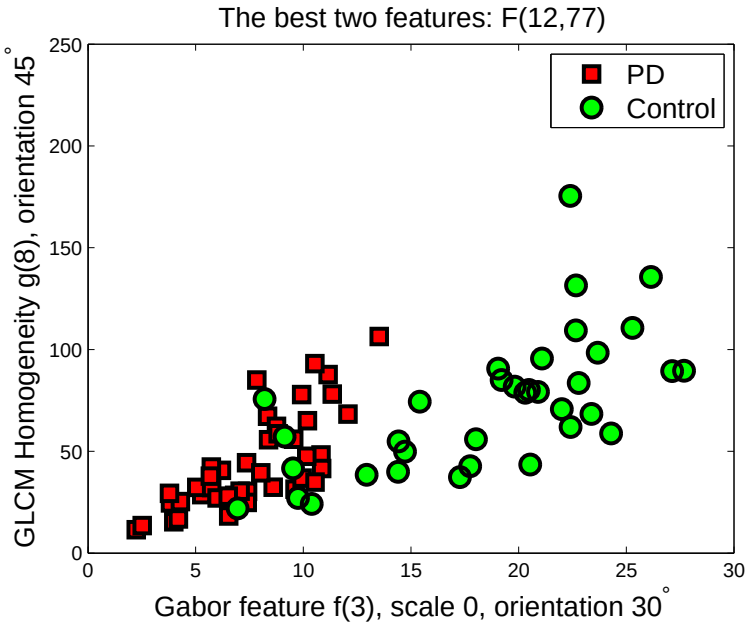


Figure 3.12: Illustration of the two selected features, Gabor features $f(3)$ and GLCM features $g(8)$.

Table 3.7: Feature selection results based on texture features, classification rates (%) of SVMs with cross-validation [38].

Method	Feature set	Accuracy	Specificity	Sensitivity
Without	F(1,...,85)	65.45	1	0
SBS	-	-	-	-
SFS	$F(73, 77)$	90.91	88.89	94.74
SFFS	$F(12, 77)$	92.73	91.67	94.74

3.4.3 Feature Selection with 101 Features

All features that were mentioned in Sections 3.4.1 and 3.4.2, in total 101 features, were evaluated using the same data. The feature selection SFS worked out an optimal feature subset that includes three features $F(12, 78, 76)$, which yield an accuracy of 93.59%. The performance of SFFS is even better than of SFS, the optimal subset consists of five features $F(12, 5, 77, 15, 11)$, which gives the best classification with 97.44% accuracy. The confusion matrix of the feature subset selected by SFFS is $\begin{pmatrix} 35 & 1 \\ 1 & 41 \end{pmatrix}$. The feature selection and classification results are shown in Table 3.8.

Table 3.8: Feature selection results based on multiple features (101 dimensions), classification rates (%) of SVMs with cross-validation. The feature subsets are selected by SFS and SFFS.

Selection	Features	Accuracy(%)
Without	F(1,...,101)	61.54
SFS	F(12,78,76)	93.59
SFFS	F(12,5,77,15,11)	97.44

3.4.4 Robust Feature Analysis

Evaluation of Rotation-invariant Gabor Filter

For a preliminary validation of the rotation-invariant Gabor filter, we chose the *University of Illinois at Urbana-Champaign* (UIUC) texture database [79] to compare two implementations of the rotation-invariant Gabor filter. The texture images were collected at UIUC by Shivani Agarwal et al. and grouped as 25 different textures, each group consists of 40 samples. We selected two different textures, T01 (bark1) and T15 (brick2) from UIUC image database, which are rotated texture sets.

We then applied three differently designed Gabor filter banks on these data, the extracted Gabor features were used for the classification by the SVM classifier. The classification was cross-validated by leave-one-out method, and the results are given in Table 3.9. The performance of the work by Zhang et al. [72] shows 76% classification rate, which is better than the method of Han et al. [71] and the conventional Gabor filter method. According to this result, we then applied the rotation-invariant Gabor filter [72] onto TCS images and compute the proposed Gabor feature, entropy, afterwards. The comparison of the conventional Gabor features and the proposed entropy feature will be given in next section.

Performance of the Proposed Gabor Feature

In this experiment, the feature vector F has 90 dimensions: $F(1, ..., 60)$ are 60 conventional Gabor features μ_{ij}, σ_{ij} in feature vector $f(1, 2, ..., 60)$; $F(61, ..., 90)$ are 30 features of entropy $f_e(1, 2, ..., 30)$. The feature vector F is listed in Table 3.10.

The normalization process was used to simulate different user settings such as brightness and contrast changes, applied to the TCS images by sonographer.

Table 3.9: The performance of each Gabor filter bank based on the public image data base, UIUC [79].

Gabor Filter	Accuracy	Confusion Matrix
Convention [66]	72	$\begin{pmatrix} 39 & 1 \\ 21 & 19 \end{pmatrix}$
Summation [71]	72	$\begin{pmatrix} 38 & 2 \\ 20 & 20 \end{pmatrix}$
Shifting [72]	76	$\begin{pmatrix} 40 & 0 \\ 19 & 21 \end{pmatrix}$

Table 3.10: The feature vector includes 60 conventional Gabor features and 30 proposed entropy features.

Feature vector	Feature name
F(1)	Conventional Gabor feature 1, $f(1)$
⋮	⋮
F(60)	Conventional Gabor feature 60, $f(60)$
F(61)	Gabor entropy feature 1, $f_e(1)$
⋮	⋮
F(90)	Gabor entropy feature 30, $f_e(30)$

In this section, three normalization methods were tested on the TCS images. The first normalization was to rescale all TCS images to full gray level range $[0, 255]$. The second method was the zero mean and unit variance ($\frac{X-\mu}{\sigma}$). Third, we applied the contrast-limited adaptive histogram equalization (CLAHE) [55] to match the histogram of ROI with a desired shape. The exponential and Rayleigh distributions were used in this experiment. As a result, the normalized data were obtained by these normalization methods. Furthermore, the Gabor features were evaluated by the feature selection SFFS method. According to the result in Table 3.11, the feature entropy is less sensitive than others features to the brightness changes of a gray level image.

The classification results were based on three datasets of TCS images, which were obtained using Philips SONOS 5500. Dataset 1 includes 36 TCS images from 21 healthy subjects and 42 TCS images from 23 PD patients. Dataset 2 includes eight control TCS images from four healthy subjects and 15 PD TCS images from 10 patients. The last dataset consists of 27 control TCS images from 14 healthy subjects and ten PD TCS images from five patients. Totally, this dataset includes 67 PD images from 38 PD patients and 71 control images

from 39 healthy subjects.

First, the Gabor filter bank was applied to the ROI of TCS images. Second, the rotation-invariant Gabor features were extracted and shifted by the dominant direction. The feature extraction was based on the manual segmentation of half of mesencephalon, marked by the physicians. Third, the feature selection was implemented on each normalized data, respectively. The criterion function was set as SVMs, the classification rates of SVMs were cross validated by leave-one-out method. The feature analysis results in Table 3.11 show that the entropy features $F(61, 77)$ (the second column) were more stable than the conventional Gabor features, mean and the standard deviation, $F(1, 5, 7)$ (the first column). The combination of features entropy $f_e(6)$ and mean $f(3)$ performed even better as shown in the third column.

Table 3.11: The performance of conventional Gabor features and the proposed entropy features based on four normalized datasets. Classification rates (%) of feature selection results by SFFS.

Normalization method	Conventional Feature F(1,5,7)	Entropy Feature F(61,77)	Optimal subset F(66,3)
$[0, 255]$	65.94	61.59	70.29
$\frac{x-\mu}{\sigma}$	67.39	60.08	70.29
Exponential	63.76	60.14	78.26
Rayleigh	30.43	68.84	77.53

Based on this data without any normalization, the features of the conventional Gabor filter in [38] achieved 69.56% accuracy. The features $F(66, 3)$ obtained by SFFS yield a better classification rate of 81.88%.

3.5 Conclusions

This chapter concentrates on selecting good combinations of features and a classifier which suits for the Parkinson's disease risk assessment based on TCS images. We proposed a hybrid feature extraction method which includes statistical, geometrical and texture features for the early PD risk assessment. The SVMs separate the input images into two classes by image characteristics other than the manual segmentation of substantia nigra. The SFS is implemented and five features including RMS contrast, Hu1-moment and other three Gabor texture features were found being the best parameters to separate control subjects from Parkin mutation carriers.

Regarding the texture analysis using Gabor filters and GLCMs for PD detection, SFFS was implemented and two features including Gabor $f(3)$ and GLCM $g(8)$ texture features were found to be the best parameters to separate control subjects from Parkin mutation carriers.

Furthermore, the classification results show that the rotation-invariant Gabor filter is better than the conventional Gabor filter. In particular, the entropy feature is more stable than the conventional Gabor features, mean and standard deviation, in the monotonic change of the gray scale.

Chapter 4

Texture Analysis of TCS Images with Scattering Operators

4.1 Scattering Convolution Network

A scattering transform computes local image descriptors by convolution with wavelet filters. With images as input, it yields scattering coefficients with modulus operators, and local averaging[45]. Similar to the local image descriptor SIFT [80], the local averaging reduces the feature variability and provides local translation invariance. Scattering coefficients are computed with a convolution network [44], they provide much richer structure information and multiscale texture variations [45]. A combined scattering representation and a combined windowed scattering estimator were introduced to preserve the information while being invariant to rotations [81].

We first apply the scattering transform for TCS images classification. Compared to the original image, the scattering image representation is much larger than the original one. As a result, the computation time for classification is large without a feature dimension reduction, even if a simple classifier is used. After obtaining the scattering coefficients, we investigate several dimensionality-reduction methods on the feature vector, the scattering coefficients, while trying to achieve rotation invariance and keep or improve the accuracy. Furthermore, we propose to use linear discriminant analysis (LDA) instead of principal component analysis (PCA) for the classification using scattering vectors. After the verification of the proposed methods based on the public image databases, we apply the dimensionality reduction methods and the classification methods by

using LDA on TCS images and demonstrate experimental results on the classification between images from PD patients and healthy controls.

4.2 Texture Analysis with Scattering Operators

Texture analysis has been used in many applications from satellite to medical images. The scattering network computes an image representation that is stable with respect to local translations and deformations.

4.2.1 Scattering Coefficients

Scattering coefficients are computed with a convolution network which cascades contractive wavelet transforms and modulus operators [82]. This way, the image information is re-transformed into co-occurrence coefficients at multiple scales and orientations [45].

Let the multi-scale directional wavelet filters be defined for any $j \in \mathbb{Z}$ and rotation $r \in G$ by

$$\psi_{2^j r}(x) = 2^{2j}\psi(2^j r^{-1}x), \tag{4.1}$$

where G is a discrete rotation group on $\mathbb{R}$ [83]. Let $\phi_J(x) = 2^{2J}\phi(2^J x)$ be a low-pass filter. The representation of an image f at a position x for scales $2^j < 2^J$ is a vector of wavelet coefficients [45]

$$W_J f(x) = \begin{pmatrix} f \star \phi_J(x) \\ f \star \psi_{2^j r}(x) \end{pmatrix}. \tag{4.2}$$

Then, the averaged wavelet coefficients are defined as [45]

$$|f \star \psi_{2^j r}| \star \phi_J(x). \tag{4.3}$$

The reason why computing the feature vectors by averaging the amplitudes of wavelet coefficients because of the investigation of Tola et al. in [84]. They observed that, more efficiently, the SIFT features can be approximately by averaging partial derivative amplitudes of f along r direction with a low-pass filter $\phi_J(x)$. Therefore, the averaged coefficients can be nearly invariant to local deformations or translations.

Furthermore, the high frequencies eliminated in (4.3) by the averaging with ϕ_J can be recovered by the convolution with another wavelet at higher layer in the

convolution network. Here, the averaged wavelet coefficients can be written as [45]

$$||f \star \psi_{2^{j_1} r_1}| \star \psi_{2^{j_2} r_2} \star \phi_J(x). \tag{4.4}$$

The scattering coefficients get their name, because they supply interference information in f for any pair of wavelets which have different scales and orientations. Coefficients are only computed for $2^{j_2} \leq 2^{j_1}$, because coefficients can be negligible at scales $2^{j_2} \geq 2^{j_1}$ [82]. According to the convolution network theory [44], this procedure can be applied q times iteratively, resulting in a vector of scattering coefficients $S_{q,J}f(x)$. A vector of scattering coefficients with m layers is then defined as [45]

$$S_J f(x) = \left(S_{q,J} f(x)\right)_{0 \leq q \leq m}. \tag{4.5}$$

From the experimental results of Mallat et al. in [82], the energy of $\|S_J f\|^2$ decays quickly as the co-occurrence order q increases. Based on the Caltech101 image database, 98% of the energy $\|S_J f\|^2$ is carried by scattering coefficients of order 0, 1, and 2 [45]. In this chapter, we therefore limit the scattering order to $m = 2$.

4.2.2 Dimensionality Reduction of the Feature Vector

Using the scattering coefficients, even though a simple classifier is used for the classification, the computation time is still considerable because of the big feature vector. Another more important question is whether each coefficient or each feature contributes to the classification rate. The vector size of the scattering coefficients for one image with N pixels is [45]

$$N_J = 2^{-2J} N \sum_{q=0}^{m} K^q \begin{pmatrix} J \\ q \end{pmatrix}, \tag{4.6}$$

where K is the number of considered directions, J is the number of scales, and m is the maximum layer of the convolution network. For example, given an input image with size of 128×128, if we choose the Morlet wavelet with four scales and six orientations, and the output of one scattering operator is a resulted 'image' with size of 16×16, after the averaging with ϕ_J and downsampling at intervals 2^J. The number of scattering operations is $1 + 24 + 216 = 241$, totally the size of the feature vector is $N_f = 16 \times 16 \times 241 = 61696$.

First, an integral approach is applied on the vector of the scattering coefficients $S_J f(x)$ to reduce the size of the feature vector. The summation of the coefficients

are calculated along K orientations at a certain scale:

$$S'_J f(x) = \alpha \sum_{r_i=1}^{r_i=K} S_J f(x), \tag{4.7}$$

where α is a weight parameter. We set $\alpha = \sqrt{K}$ in the experiment. Now the scattering coefficients $S'_J f(x)$ are invariant to rotations. We call this method the integral approach I. In addition, the size of the feature vector N_f is reduced to 10496 with $K = 6$.

Furthermore, all the scattering coefficients at the same convolution layer are integrated together (the integral approach II):

$$S'_J f(x) = \beta \sum_{q=1}^{m} S_J f(x), \tag{4.8}$$

where β is inversely related to the numbers of the scattering coefficients at each layer, for example $\beta = (\frac{1}{24}, \frac{1}{216})$. The size of the feature vector is $N_f = 16 \times 16 \times 3 = 768$.

In this approach, similar to the conventional Gabor feature 'mean', a scattering representation $S_J f(x)$ is approximated with the expected value of each resulted image:

$$S'_J f(x) = \left(E\{S_{q,J} f(x)\}\right)_{0 \le q \le m}. \tag{4.9}$$

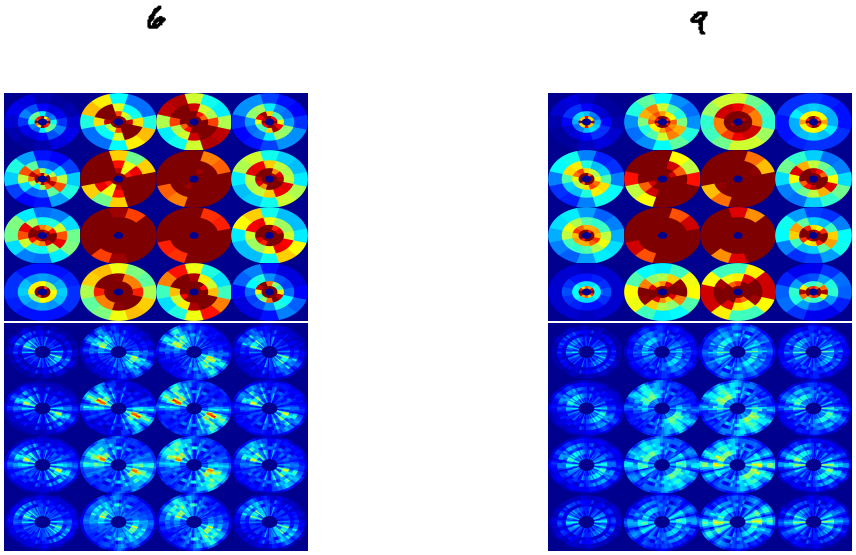
In this case, the size of the feature vector N_f is reduced from 61696 to 241 with $K = 6$ and $J = 4$. The method will be referred to as 'mean' in the following. The size of scattering features and the reduced feature vectors based on the MNIST database and the CUReT texture database are given in Table 4.1. The size of one resulted image for them are 4×4 and 32×32, respectively.

Table 4.1: The feature vectors computed by the dimensionality reduction methods. The figures between the brackets are the number of feature subset in each convolution layer.

	Feature Subsets	Scattering features	
		MNIST	CUReT
Original	[1][24][216]	3,856	162,916
Integral I	[1][4][36]	656	41,984
Integral II	[1][1][1]	48	3,072
'mean'	[1][24][216]	241	241
Resulted image		4×4	32×32

4.2.3 Verification of Feature Reduction Methods

The proposed dimensionality reduction methods were performed on the MNIST database and the CUReT texture database. The MNIST database of handwritten digits is a good database for learning and testing pattern recognition methods. It consists of a training set of 60,000 training samples and a test set of 10,000 samples. An example of digit '6' and '9' with the corresponding normalized scattering vectors are shown in Figure 4.1. The scattering feature are displayed as 4-by-4 array at convolution layer 1 and 2, as shown in the second and third row of Figure 4.1, respectively.



(a) (b)

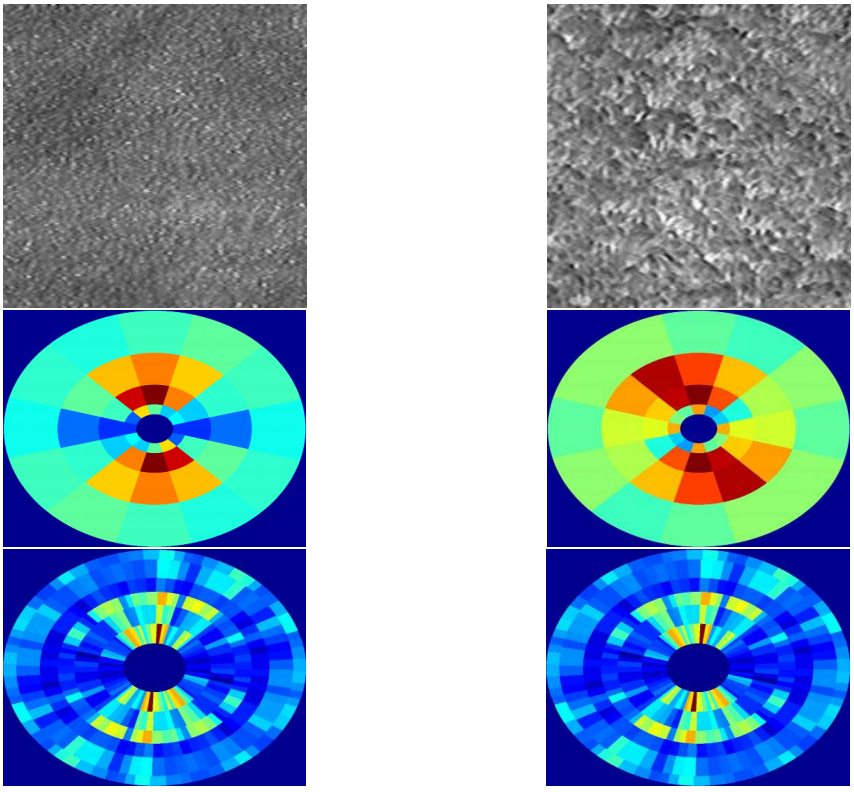
Figure 4.1: The images of digits (a) '6' and (b) '9', and the corresponding scattering vectors. The second and bottom row contains arrays of scattering vectors for $m = 1$ and $m = 2$, respectively.

For the validation of the feature vector reduction methods, the computations are implemented on Scattering toolboxes for MATLAB [85] with the PCA model at the maximum orders $m = 2$. Using the PCA affine selection model, the performance of the original coefficients vector $S_J f$ with the two integral vectors $S'_J f$ are compared. Three different sizes of training sets are tested, and the percentages of errors are shown in Table 4.2. The sizes of the feature vectors reduced by PCA with different training sets are 83, 44, and 133, respectively.

Table 4.2: Percentage of errors (%) on the MNIST database. The 'Integral I' and 'Integral II' approaches are shown in 4.7 and 4.8, respectively.

Training	PCA	Integral 1	Integral 2	mean
200	1.46	1.89	13.23	5.41
300	1.54	1.68	11.71	5.05
2000	0.78	1.19	11.00	3.58

Texture image classification with scattering operators is tested on the Columbia-Utrecht reflectance and texture database (CUReT)[86]. The CURet includes 61 classes of image textures of $N = 200^2$ pixels, and each texture class with the same material but over different combinations of viewing and illumination conditions [45]. Two examples and the corresponding scattering vectors are shown in Figure 4.2.



(a) (b)

Figure 4.2: Texture images from class (a) 'sample01' and (b) 'sample03' in the CURet database and their scattering vectors for $m = 1, 2$.

The database is separated into a training and a test set randomly. The training set consists of 23 or 46 images and the results are averaged over ten different

groups. The percentages of errors with the original scattering feature vector and the reduced features based on the CURet database are listed in Table 4.3.

Table 4.3: Percentage of errors (%) on the CUReT database.

Training	PCA	Integral 1	Integral 2	mean
23	1.67	0.97	1.67	1.25
46	0.72	0.36	0.43	1.30

The first integral method achieved better accuracy than the whole scattering vector on CUReT database due to the rotation invariance as shown in Table 4.3. The possible reason that for the MNIST database the first integral method performs worse than the full scattering vector may be the fact that both digits '9' and '6' are included in the MNIST database, which are indistinguishable with a rotation invariant feature set. As the Tables 4.2 and 4.3 show, the second integral method, which produces the shortest feature vector, cannot yield a classification rate as good as the original feature set.

4.3 Classification Using Scattering Vectors

Based on the scattering transform, a distance between two images f and g can be defined as [45]

$$\|S_J f - S_J g\|^2 = \int |S_J f(x) - S_J g(x)|^2 dx, \tag{4.10}$$

where $|S_J f(x)|^2$ is the squared Euclidean norm of the coefficients $S_J f(x)$.

Classifier based on PCA Affine Space Selection

Let $S_J F_c$ be the family of the coefficient vectors in the class c, and $E\{S_J F_c\}$ be their expected values. The difference between them is given as

$$X_c = S_J F_c - E\{S_J F_c\}. \tag{4.11}$$

X_c can be represented by its projection $V_{d,c}$ computed by principal component analysis (PCA) onto a lower dimensional ($d \ll N_J$) linear space [82]. The dimension d is obtained by cross validation in the training stage [45].

A simple generative classifier based on the scattering distance was introduced in [83]. It approximates the scattering distance from a sample image $I_t(x)$ to

class centers by affine-space models computed with PCA. Let $W_{d,c}$ be the weight matrix of PCA for the class c with the d biggest eigenvalues, the classifier based on the PCA affine space selection is

$$\hat{c}(I_t) = \underset{c \leq n}{\operatorname{argmin}} \left(\| X_t \|_2^2 - \| (X_t)^T \cdot W_{d,c} \|_2^2 \right), \tag{4.12}$$

$$X_t = S_J I_t - E\{S_J F_c\}, \tag{4.13}$$

where n denotes the number of classes, $S_J I_t$ are the scattering coefficients of a test image I_t.

4.3.1 Classification by LDA

In general, the data projected by PCA approximates the original data in a least-square sense. The first principal component, the eigenvector with the largest eigenvalue of the covariance matrix, is corresponding to the direction with the largest variance of data. The objective of LDA is to project the data onto a space in which one can maximize the separation between data in different classes. Compared to PCA, LDA projects and separates the data in a least-squares sense.

Assume a set of m-dimensional samples $x = [x_1, x_2, ..., x_m]^T$ of which n_1 belong to class ω_1 and n_2 belong to class ω_2. We seek to calculate a linear combination of the components of x using a weight matrix $W = [w_1, w_2, ..., w_m]^T$, to project the samples x onto a space (a line in a two-classes case) in which one can maximize the distance between the centers of classes. In other words, a scalar y is obtained by the dot product

$$y = W^T x. \tag{4.14}$$

The mean and the scatter matrix are given as

$$\mu_i = \frac{1}{n_i} \sum_{x \in w_i} x, \quad S_i = \sum_{x \in w_i} (x - \mu_i)(x - \mu_i)^T, \tag{4.15}$$

respectively. The Fisher linear discriminant takes into account the standard deviation within the classes, the distance between the projected means is normalized by a measure of the within-class variability, the criterion function is defined as [87]

$$J(W) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2}. \tag{4.16}$$

For the case of two classes, the within-class scatter and the between-class scatter are computed as $S_W = S_1 + S_2$ and $S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$, respectively. We differentiate $J(W)$ with respect to W to obtain the maximum of $J(W)$, the solution is the eigenvector with the biggest eigenvalue as given [87]

$$W^* = \underset{W}{\operatorname{argmax}} J(W) = \underset{W}{\operatorname{argmax}} \left(\frac{W^T S_B W}{W^T S_W W} \right) = S_W^{-1} S_B. \tag{4.17}$$

For the c-class problem, multiple discriminant analysis, the within-class scatter and the between-class scatter are computed as [87]

$$S_W = \sum_{i=1}^{C} S_i, \quad S_B = \sum_{i=1}^{C} n_i (\mu_i - \mu)(\mu_i - \mu)^T, \tag{4.18}$$

respectively, where μ is mean of the total samples x. As in the two-class case, for c classes we still seek a weight matrix W that maximizes the ratio of the between-class scatter to the within-class scatter, a simple measure is the determinant of the scatter matrix [87]. The criterion function for c-class problem is given by this measure [87]:

$$J(W) = \frac{|\tilde{S}_B|}{|\tilde{S}_W|} = \frac{|W^T S_B W|}{|W^T S_W W|}. \tag{4.19}$$

The simple solution to maximize $J(W)$ is to find an optimal W that consists of the generalized eigenvectors with the largest eigenvalues in [87]

$$S_B w_i = \lambda_i S_W w_i. \tag{4.20}$$

In order to avoid the case in which S_W is singular and the unnecessary computation for the inverse of S_W, one can obtain the eigenvectors by solving [87]

$$(S_B - \lambda_i S_W) w_i = 0. \tag{4.21}$$

Because S_B is of rank $c-1$ or less the weight matrix W consists of $c-1$ projection vectors

$$W^* = [w_1^* w_2^* ... w_{c-1}^*]. \tag{4.22}$$

Thus, the weight matrix W can be used to project the original data from d dimensions onto $c-1$ dimensions. Then, a Bayes classifier is designed for finding a threshold separating the projected data.

In practice, the computation problem would arise that the scatter matrices are difficult to obtain when the observations are much less than the features. Here, three solutions are presented to apply LDA on the scattering features. As two

examples shown in Figure 4.3, that are the digits '6' and '9' taken from the MNIST database. The scattering features obtained from each sample with size 3856 (16×241) are displayed in 2D. The resulting images with size of 4×4 pixels obtained from each scattering operator is rearranged as a vertical line, in total of there are 241 operators. As a simple solution, LDA can be applied onto the scattering features vertically, line by line, as the illustration in Figure 4.3. It starts from the evaluation of the feature subset corresponding to each operator, and a selection can be followed afterwards. The method will be referred to as 'operator selection' in the following. Similarly, LDA can be implemented horizontally on scattering features that can evaluate the performance of each pixel along all resulted images. This method will be referred to as 'pixel selection' in the following.

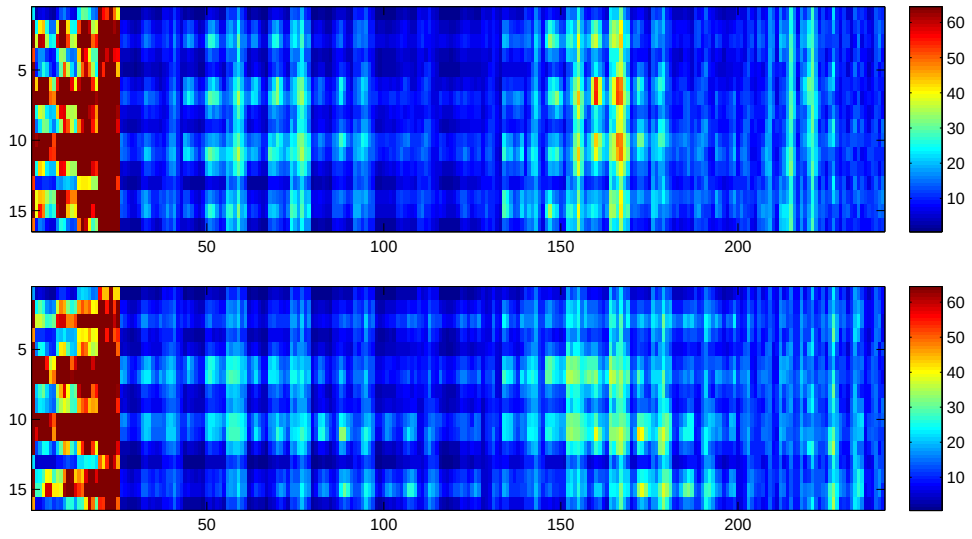


Figure 4.3: The scattering coefficients of the digit '6' (upper) and '9' (bottom) are displayed in 2D. The output of each operator is rearranged as a vertical line with 16 pixels, in total of 241 scattering operators.

Evaluation on Each Operator

First, LDA is applied on each feature subset that contains coefficients obtained by the scattering operators to evaluate the individual performance. Assume y_i is a feature subset of $S_J f(x)$ that consists of 16 features, $i = 1, 2, ..., 241$. Let $\hat{y}$ be the classification results based on 241 scattering feature subsets, each classifier votes for one class $\hat{y}_i$. Based on $\hat{y}$, the final classification of one test sample is

done by the maximum voting

$$\hat{c}(I_t) = \underset{c \leq n}{\operatorname{argmax}}(P_c), \tag{4.23}$$

where P_c is the probability of class c, in other words, how many classifiers classify I_t as class c in $\hat{y}$.

The performance of each feature subset using a LDA-based classifier with 300 training samples in the MNIST database is shown in Figure 4.4 a. The features are selected sequentially by adding the local best feature subset, which provides the lowest error rate, to a candidate set. As shown in Figure 4.4 b, the candidate set achieved the lowest error rate with 64 feature subsets. That means only 64 selected scattering operators instead of the total 241 operators can be implemented on the test set.

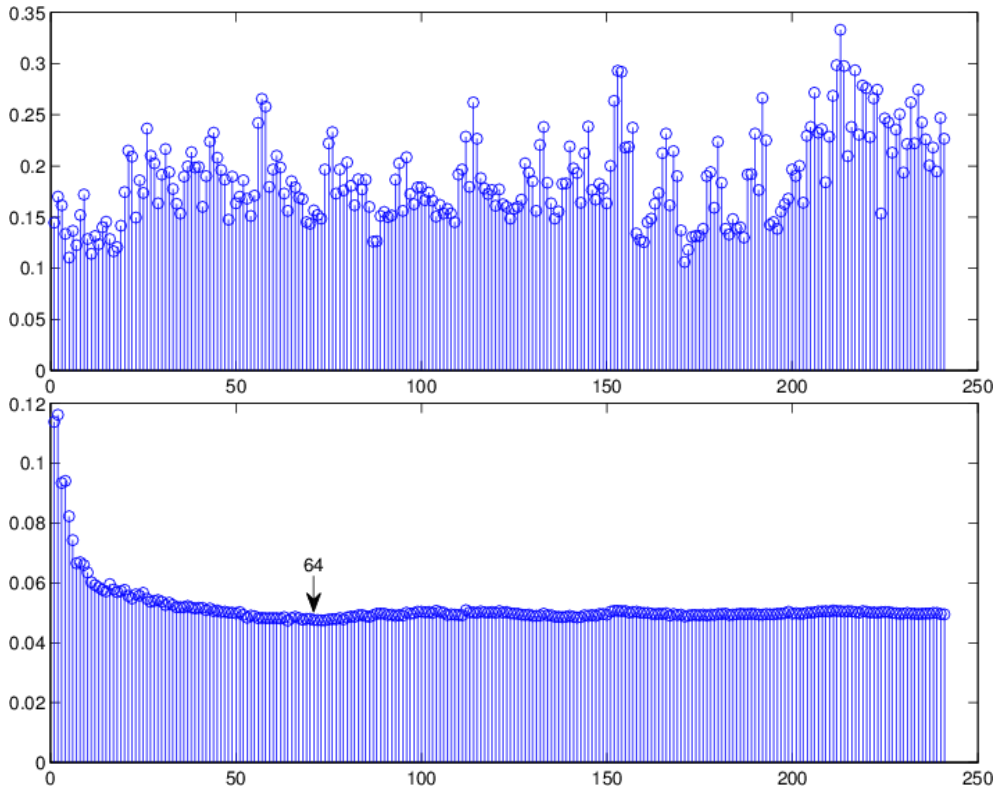


Figure 4.4: The performance of each feature subset and the operator selection. The error rate (vertical axis) of each feature subset by LDA with a training set of 300 samples (upper), in total of 241 classifications (horizontal axis). The lowest error rate can be achieved by 64 selected feature subsets (bottom).

Evaluation of the Pixel Selection

Second, LDA is applied on the feature subset that contains the same pixel along all obtained images. In this case, assume y_j contains the values of the same pixel in the output of each operator, $j = 1, 2, ..., 16$. The similar procedure as in the previous section is implemented on each feature subset y_j. The performance of each feature subset with 300 training samples is shown in Figure 4.5 a. After the pixel selection, the selected features that are corresponding to a ROI consist of seven pixels that yield the best classification as shown in Figure 4.5 b.

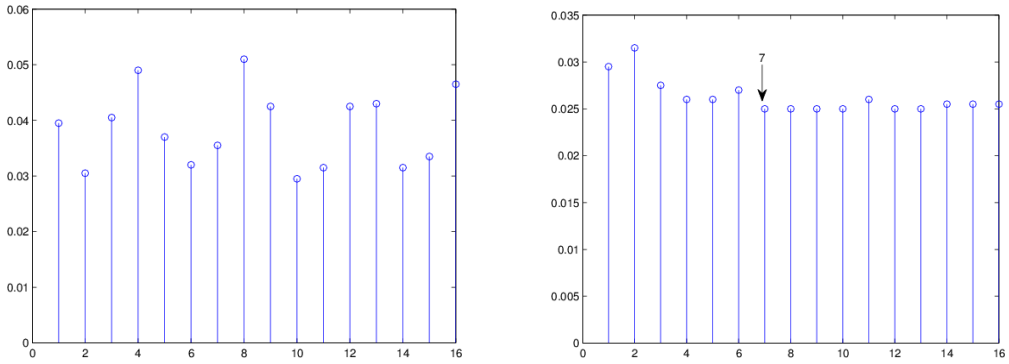


Figure 4.5: The performance of the pixel selection. (a) The error rate (vertical axis) of y_j by LDA with training set of 300 samples, in total of 16 pixels in each resulted image (horizontal axis). (b) The illustration of the pixel selection, the lowest error rate is achieved by seven feature subset that are corresponding to seven pixels.

4.3.2 Feature Selection for LDA

One problem using scattering coefficients for the classification is that the features $S_J f(x)$ are much more than the samples. For example, with a training set of 300 samples in the MNIST database, the size of the scattering coefficients for one sample is 3856 ($3856 \gg 300$). In this section, the strategy is to select features from $S_J f(x)$ that can achieve better perfromance of the LDA-based classifier. According to the criterion function of LDA(4.16), the features with bigger distances between the centers of classes, while with smaller variances, might support the maximization of $J(\cdot)$. Let $\hat{F}$ be the selected features

$$\hat{F} = \underset{l \leq L}{\operatorname{argmin}} \left(\frac{\sum_{i=1}^{C} \sigma_{l,i}^2}{\sum_{i=1}^{C} n_{l,i} |\mu_{l,i} - \mu_l|^2} \right), \tag{4.24}$$

where L indicates the size of the feature vector, C is the number of classes, and $\sigma_{l,i}$ is the variance of class i for the feature l.

For instance, given a 3-class training set with 300 samples of each class that were randomly selected from the MNIST database. Then, the weight matrix W^T of LDA is obtained based on 300 features that are randomly selected from 3856 features. The weight matrix W^T consists of $c-1$ projection vectors. The data projected by the first two vectors with the biggest eigenvalues are displayed in Figures 4.6, the upper row. In addition, 50 features were selected by using the criterion function(4.24). The data projected by the weight matrix computed with the selected features is shown in Figure 4.6, the bottom row. Apparently, the projection vectors were computed based on the selected features provided higher discrimination power than the features without selection.

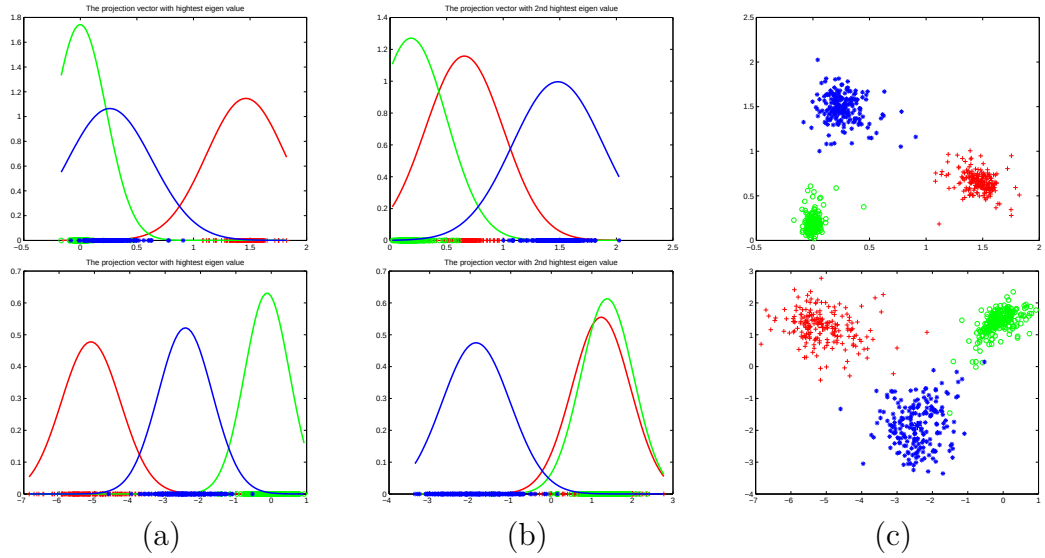


Figure 4.6: The illustration of the data projected by the first two weight vectors (a) w_1 and (b)w_2. (c) 2D display of the projected data. The weight matrices were computed with 300 features (top row) and 50 selected features by equation 4.24 (bottom row), respectively.

In addition to the equation(4.24), other criterion functions can also be used to select features for LDA, such as the total distance of classes $d_M = \sum_{i=1}^{C} n_i(\mu_i - \mu)^2$ or the total variance of classes $v_M = \sum_{i=1}^{C} \sigma_i^2$. The data projected by the two weight vectors were displayed in Figure 4.7, the weight matrices were computed based on the selected features by v_M and d_M, respectively.

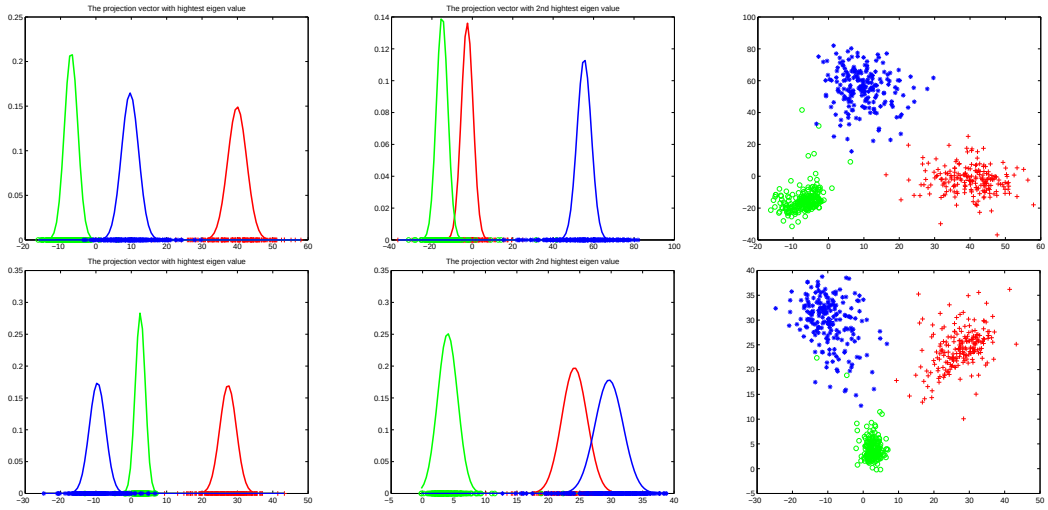


Figure 4.7: The illustration of the data projected by the first two projection vectors (a) w_1 and (b) w_2. (c) 2D display of the projected data. The weight matrices were computed based on 50 features selected by the criteria v_M (top row) and d_M (bottom row), respectively.

4.3.3 Verification of Classification by LDA

Results for the Selected Operator

The LDA-based classifier was tested based on the training sets with 300, 1000, and 2000 samples in the MNIST database. The results for the maximum voting output among all 241 feature subsets were given in the first column in Table 4.4. The classification results based on the selected feature subsets were better than using all features. For example, with 300 training samples, 64 feature subsets were successively selected that yield the error rate of 5.18%. It is lower than 5.24% with all 241 feature subsets. As a result, only the corresponding operators that yield the selected feature subsets need to be implemented for the test set. That might reduce almost 3/4 computation involving the convolutions and classifications, while keeping or improving the accuracy.

Furthermore, the selected feature subsets were combined into one feature vector and applied with PCA to reduce the dimensionality. The best dimension d was obtained by cross-validation during the training. The results based on one LDA-based classifier were given in the third column. As example with 300 training samples, 64 feature subsets were firstly selected and then composed of one feature vector, the size is 1024 (16×64). By using PCA the data was mapped

Table 4.4: Percentage of errors (%) based on all 241 feature subsets and the selected feature subsets on the MNIST database. The performance of the selected features with PCA was the best and shown in the third column. 'd' indicates the obtained best dimension.

Training size	All 241 feature subsets	Selected feature subsets	Selection and PCA	Affine model selection [45]
300	5.24	5.18(64)	2.47(50 d)	6.05
1000	4.98	4.81(86)	1.90(80 d)	2.39
2000	4.90	4.66(98)	1.60(130 d)	1.71

from 1024 onto 50 dimensions, on which the best classification was achieved during the training. Compared with the previous work (Mallat et al.) in 2011 [45] (the last column), the proposed method achieved better performance.

Results for Selected Pixels (ROI)

In order to find out which pixels or ROI provide higher discrimination power, the scattering coefficients were grouped into 16 feature vectors by the pixel selection approach, each feature vector consists of 241 features. The classification with the maximum voting among 16 feature vectors were shown in the first column, referred as 'All pixels' in Table 4.5. The results for the selected feature subsets were better than those without selection. For example, with 1000 training samples, the selected feature subsets corresponding seven pixels (ROI) contributed 2.72% error rate, which is lower than 2.83%.

Table 4.5: Percentage of errors (%) of the scattering coefficients grouped by the pixel selection approach on the MNIST database.

Training	All pixels	Selected pixels	Size of ROI
1000	2.83	2.72	7
2000	1.98	1.91	11

Furthermore, discrete cosine transform (DCT) is applied on scattering coefficients and the original images, respectively. The results are given on the first, third row in Table 4.6. The second and forth row list the results for the selection method, the figures between parentheses are the number of the selected pixles or the size of ROI. Mallat et al. found that applying DCT on scattering coefficients cannot affect any linear classifier in [82]. Our experiment results approved their

finding. For example, the results for the feature subsets with or without the selection were quite similar as listed in the first two columns.

Table 4.6: The experiments with DCT. 'DCT Coeff' and 'DCT Image' indicate that applying DCT on scattering coefficient and the original image, respectively. 'DCT4' means 2×2 block-wise DCT.

Training	DCT Coeff	DCT4 Coeff	DCT Image	DCT4 Image
1000	3.11	2.92	5.30	4.36
	2.97(11)	2.88(13)	4.66(3)	4.05(5)
2000	2.34	2.04	3.79	3.63
	2.24(3)	2.04(16)	1.98(11)	2.47(7)

Moreover, as the previous work in last section, PCA is applied on the 16 feature vectors grouped by the pixel. The classification results and the best dimension d are listed in the second column in Table 4.7. The error rates are similar to the results obtained by the operator selection as listed in the third column in Table 4.4. The classification given by 2×2 block-wise DCT on coefficient (denoted as DCT4 Coeff) were better than others.

Table 4.7: The feature vectors were reduced by PCA for classification. The figures between parentheses are the best dimension d.

Training	Original	DCT Coeff	DCT4 Coeff	DCT Image	DCT4 Image
300	2.37(50)	2.24(40)	2.28(50)	2.79(50)	3.75(50)
1000	2.04(80)	1.95(60)	1.87(60)	2.21(80)	3.06(100)
2000	1.67(110)	1.70(130)	1.57(60)	2.06(70)	2.79(110)

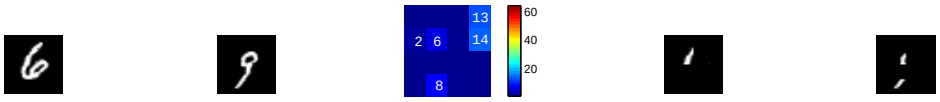
In addition, PCA was firstly applied on 16 feature vectors individually on the dimension d, and the pixel selection was followed afterwards. The results in Table 4.8 were even better than the previous works in this section. For example, based on 2000 samples the error rate 1.65% was achieved by the selected features without DCT, compared with the error rate 1.67% and 1.91% in Table 4.7 and Table 4.5, respectively.

More interestingly, the pixel selection is applied on the data that only contains the digit '6' and '9'. Based on 1000 training samples, the LDA-based classifier achieved 0.05% error rate with all 16 feature vector grouped by the pixel. The selection method worked out the ROI consists of pixel number $2, 6, 8, 13, 14$ as shown in Figure 4.8 c. The corresponding feature subsets could achieve perfect

Table 4.8: The feature vectors were applied by PCA before and the selection followed afterwards.

Training	Original	DCT Coeff	DCT4 Coeff	DCT Image	DCT4 Image
1000	2.04(16)	1.85(12)	1.78(15)	2.19(7)	2.87(5)
2000	1.65(7)	1.68(7)	1.52(13)	1.98(11)	2.47(7)

classification with zero error. These areas can provide higher discrimination power than other area on separating the classes '6' and '9'.



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

Figure 4.8: The ROI has dominant contribution for separating the digit (a) '6' and (b) '9'. (c) The ROI selected by the pixel selection. (d) and (e) show the selected pixels from (a) and (b), respectively.

Results for Selected Features

This experiment was designed to evaluate the performance of the selected features by the criterion function for LDA-based classifier. First, the criterion values were calculated based on the scattering features, such as the proposed selection criterion function(4.24), i.e., the total distance of classes. Second, features were selected sequentially by adding the local best feature which provides the lowest criterion value to a candidate set. At last, the dimension d can be obtained by the cross-validation in the training stage. For example, 64 features with the biggest values were selected with 200 training samples, the classification was achieved 6.15% based on this feature subset.

Considering that the big difference of distances between each two classes might affect the calculation, we introduced a new parameter τ into the criterion function:

$$\tau = \frac{\underset{i,j \leq C}{\operatorname{argmin}} |\mu_i - \mu_j|^2}{\underset{i,j \leq C}{\operatorname{argmax}} |\mu_i - \mu_j|^2}. \tag{4.25}$$

The results for the selected features are listed in Table 4.9. From our experiments, we found the perfromance of the distance of classes d_M was more stable

and better than others as shown in the last column.

Table 4.9: The results for the the selected features by the criteria based on LDA. The figures between parentheses are the best dimension d obtained by cross-validation.

Training	$\frac{v_M}{d_M}$ 4.24	$\frac{v_M}{\tau d_M}$	d_M
200	6.15(64)	5.50(64)	5.04(50)
300	5.06(77)	3.90(73)	3.43(66)
1000	3.27(221)	2.77(113)	2.44(80)
2000	2.61(191)	2.50(133)	2.04(110)

Furthermore, the scattering features were reordered with respect to the criterion d_M. Based on the results in Table 4.9, the ranked feature vectors were splited into separate segments, the size of each segment was set as 80 elements. In other words, the features are grouped based on the criterion d_M, then implemented classification on individual feature subsets. The final decision was made by the maximum voting among these classifications.

For example, the size of the scattering features for one sample is 3856, with one subset containing 80 features, the total number of the classification is around $3856/80 \approx 38$. The results including the experiments with DCT are given in Table 4.10.

Compared with the previous works [45, 82], with the selection, the LDA-based classifier achieved better classification than the original works based on the PCA affine selection model [45, 82]. But with the training set increasing, the latest work [82] can yield even lower error rate. For example, with 2000 training samples, the percentage of accuracy 98.7% in [82] was marginally better than the proposed method with 98.4%.

Table 4.10: The results for the feature selection based on LDA. Classification based on 38 feature subsets, each subset includes 80 ordered features. The criterion function is the distance of classes d_M.

Training	Original	DCT Coeff	DCT Image	Mallat[45]	Mallat[82]
300	2.40	2.27	3.79	6.05	4.70
1000	1.95	2.07	2.95	2.39	2.30
2000	1.61	1.59	2.71	1.71	1.30

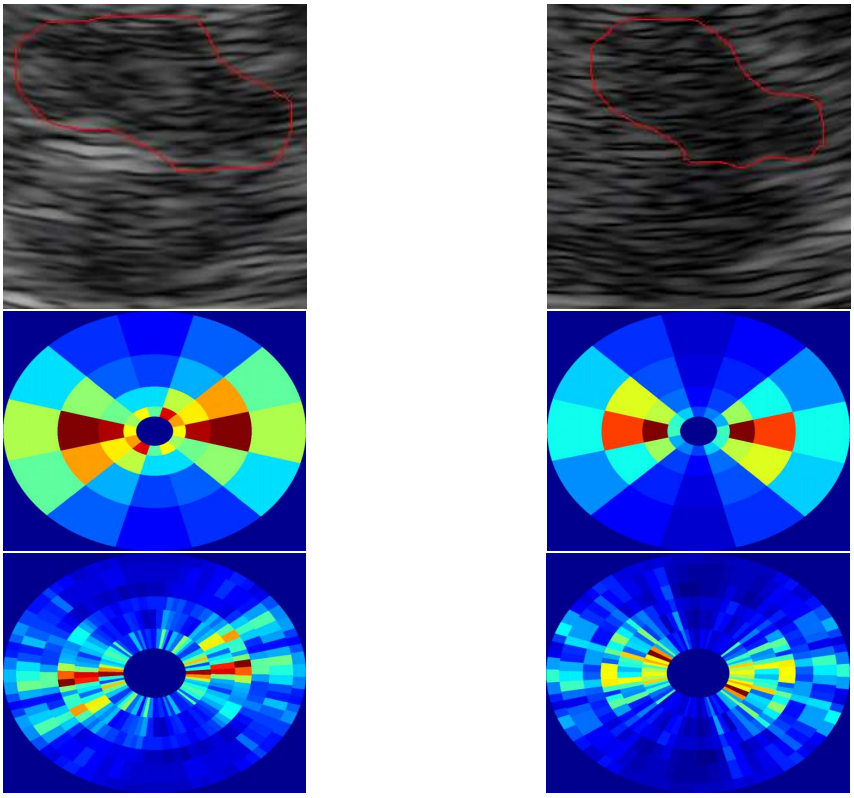
4.4 Experimental Results for TCS Images Classification

The experiments of classification with scattering coefficients were implemented with TCS image data. The illustration of two TCS images and the corresponding scattering coefficients are shown in Figure 4.9.

The experimental materials include two datasets with the structure of four groups. The first group consists of 21 patients with both PD and Parkin mutation; Group 2 consists of 12 patients with the Parkin mutation; Group 3 includes 16 patients who were diagnosed as PD patients without the Parkin mutation; The last group only contains TCS images from healthy control subjects without the Parkin mutation. Here, a study has been conducted to evaluate whether these scattering features can be used as a PD indicator and/or a genetic mutation indicator. First, we combined the images of group 1 and group 3 as PD data for the comparison with the healthy controls (group 4). Second, the TCS images in groups 1 and 2 were combined as Parkin mutation data to test whether the local features can separate Parkin mutation carriers from the healthy controls (group 4).

The scattering representations of TCS images were computed with a Morlet wavelet. For testing the dimensionality reduction methods, the classification results were obtained with the PCA affine model classifier [82] and support vector machine (SVM) classifier with the sequential minimal optimization (SMO) method. For the evaluation of the feature selection methods, the classifications were implemented by the LDA-based classifier. In addition, the pre-processing of the input images was necessary and introduced as following.

In the previous work [32], from the prior knowledge of the anatomic location of HoM and SN, the mask was created from the ellipse which was fitted onto the ROI as mentioned in [27]. This mask was used for the exclusion of the detected blobs which are outside of the HoM region. Here, in order to align the HoM region of each subject for the same orientation, each manually segmented boundary of HoM is fitted with an ellipse and then transformed onto a target ellipse with the affine transformation. The centers of the ellipses and the eight control points on the ellipses are applied to calculate the transition matrix of the affine transform. The target ellipse and one fitted ellipse with a ROI are shown in Figure 2.7. The center point and eight control points are indicated in Figures 2.7 (a) and (b). Using the transition matrix, the HoM and SN areas of the manual segmentations are transformed onto the target ellipse (Figures 2.8



(a) (b)

Figure 4.9: A TCS image from (a) a PD patient and (b) a healthy control subject (Philips SONOS 5500). The scattering coefficients for $m = 1$ and $m = 2$ were shown on the second and third row, respectively.

(a) and (c)). The original image of the HoM region and its affine adaptation result can be seen in Figures 2.8 (b) and (d), respectively.

4.4.1 Performance of Feature Reduction Methods

The results with the scattering coefficients for the TCS image classification were obtained by the PCA affine model classifier [82] and are shown in Tables 4.11 and 4.12. First, the rotation normalization worked better on Parkin mutation and control data than on other data. Second, the first integral method did not achieve better performance than the original feature vector.

According to the classification rates in Tables 4.11 and 4.12, the scattering coefficients did not supply better accuracy than the previous works [32, 39]. To improve the performance, the forward sequential feature selection (SFS)

Table 4.11: Percentage of errors (%) on the original and normalized TCS data: PD and control data.

Data	PCA	Integral I	mean
Original	27.52	32.11	26.61
Normalized	28.44	33.03	33.03

Table 4.12: Percentage of errors (%) on the original and normalized TCS data: Parkin mutation and control data.

Data	PCA	Integral I	mean
Original	31.68	30.99	25.74
Normalized	26.73	26.73	38.61

method was used on the scattering 'mean' features. The criterion function of SFS was SVMs with the leave-one-out cross validation method. The input of SVM classifier were the mean values of the coefficient sets $S_{q,J}f(x)$. The Gaussian radial basis functions (RBF) were selected as the kernel function. The classification results of TCS images based on the selected features are shown in Table 4.13.

Table 4.13: Percentage of errors (%) with the selection method on the normalized TCS data: PD and control data (data 1), Parkin mutation and control data (data 2). The scattering coefficients used as the 'mean' features.

Data	Feature subset	Error
data 1	$F(93, 5, 133, 22, 1, 82)$	11.01
data 2	$F(130, 71, 188, 11, 64, 110, 4, 111)$	9.90

The feature vector $F(1, 2, ..., 240, 241)$ consists of mean values of the coefficient sets $S_{q,J}f(x)$, SFS selected six features that are corresponding to six coefficients based on the PD and control data. The error decreased from 28.44% to 11.01%. In another data subset, Parkin mutation and control data, the selected feature subset includes seven features (coefficients) and decreased the error rate from 26.73% to only 9.90%. The selected scattering coefficients provide better classification rates than the previous works [32, 39] on the large datasets.

4.4.2 Classification Results Based on LDA

Based on the verification on the MNIST database, three classification methods are applied on TCS images based on the LDA-based classifier, which are the scattering operator selection, the pixel selection, and the feature selection. The first step was to compute the scattering coefficients with the TCS images from PD and healthy control subjects. In the second step, three selection methods are implemented based on the obtained scattering features. Particularly, the training set of TCS data is much smaller than the MNIST database. We therefore applied PCA on the input features for the dimensionality reduction in order to compute the weight matrix for LDA in each classification. The final decision was obtained by the maximum voting procedure based on classifications.

The results for separating PD from the control were given in Table 4.14. The performance of the scattering operator and pixel selection methods were quite similar, the error rate was 12.5%. For the feature selection by the criterion function, the results were calculated with the feature subset which size of 200 features. In this experiment, the criterion was chosen as the distance of classes d_M.

Table 4.14: Percentage of errors (%) on the original and normalized TCS data: PD and control data.

	Coefficient selection	Pixel selection	Feature selection
Without selection	18.27%	14.42%	20.19%
Selection	12.5%	12.5%	18.27%
Number of All	241	256	205
Number of Selected	33	60	30

4.5 Conclusions

A scattering transform provides co-occurrence coefficients that contain more information than the conventional Gabor filters. The scattering coefficients obtained in the first convolution layer are similar to the Gabor filter bank that was mentioned in Chapter 3. But the size of the scattering features is much larger than for the Gabor features. Although the complexity of the PCA affine selection classifier is lower than for SVMs, the computations for the convolutions and the classifications are really heavy. Hence, we then implemented the methods for the dimensionality reduction and the classification approaches with

LDA-based classifier to avoid heavy computation using the affine space selection. With the integral approach, the size of the scattering vector was reduced and rotation invariance was achieved. With the selected scattering features, the proposed classification methods based on LDA-based classifier were much faster than the PCA classifier while resulting in a similar accuracy.

After the verification by the two image databases, the scattering operators were applied on TCS images. Based on the integral approach, with the feature selection method SFS, the selected coefficients performed significantly better than the whole scattering-coefficient vector. The similar performance was achieved by the proposed classification methods based on LDA.

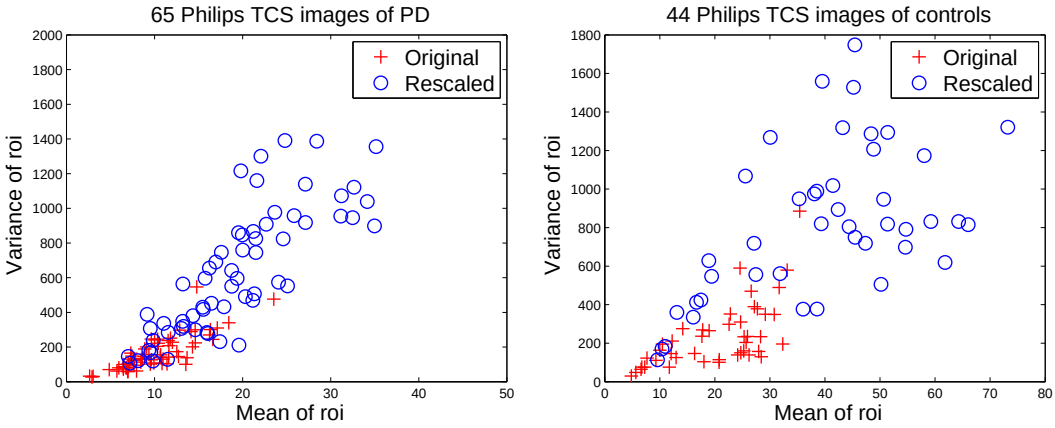
Chapter 5

Local Feature Analysis for Hyperechogenicity Estimation

5.1 Motivation of Local Feature Analysis

In this chapter, a large experimental data is analyzed while taking into account the affection from the different settings on the properties of the TCS images, such as the gray values, the brightness, and the contrast. The considerable variability among different datasets was illustrated in the previous work [32]. Our goal is to develop local features that are invariant to the illumination and contrast changes. The proposed local feature analysis method applies invariant blob detection to localize the hyperechogenicity area in ROI and extracts local features based on watershed regions for the hyperechogenicity estimation. Compared to the statistical and texture features, the local features could provide the visual characters of ROI for neurologist, such as the location and area of the hyperechogenicity.

In order to study the influence of a brightness and contrast changes, three normalization methods are applied to TCS images to simulate the different user settings. First, for each image the intensity values were rescaled to the range $[0, 255]$ as given by $I_{new} = (I - I_{min})/(I_{max} - I_{min}) \cdot 255$, where I_{min} and I_{max} are the minimum and maximum gray values, respectively. A considerable difference between the original and the normalized images is illustrated in Figure 5.1. Second, each image is normalized to zero mean and unit variance by $I_{new} = (I - \mu)/\sigma$, where μ and σ are the mean and the standard deviation, respectively. Third, the contrast-limited adaptive histogram equalization(CLAHE) [55] is used to



(a) (b)

Figure 5.1: The illustration of the mean and variance of ROI (half of mesencephalon) of 109 TCS images. (a) 33 subjects of Parkinson's Disease. (b) 25 subjects of healthy control. Red plus and blue circle indicates the original images and the rescaled images, respectively. (the intensity values are rescaled to $[0, 255]$).

match the histogram of the image with a desired shape such as the Rayleigh and exponential distributions.

5.2 Invariant Scale Blob Detection

The first automatic SN echogenicity analysis in 3D TCS based on random-forest was proposed in 2012 [43]. In their method, the volumetric SN echogenicity detection depends on the quality of the reconstructed volume from the obtained video sequences of the 2D TCS images. In this section, we focus on the robust image analysis method for the SN echogenicity detection from 2D TCS images. Therefore the regions of ROI segmented manually by physicians and then are analyzed for the estimation of the hyperechogenicity. In order to extract the features of TCS images that are not sensitive to this influence from the different settings, a local feature analysis method is proposed which applies invariant blob detection to ROI and extracts local features for the hyperechogenicity estimation.

In a clinical examination of PD using TCS, the size of the hyperechogenicity in SN area is used for the diagnosis.The hyperechogenicity of SN area consists of several bright spots in TCS image. The blob detection algorithm is stable

under the monotonic changes in gray scale. The goal of this section is to localize the hyperechogenicity in ROI by the invariant scale blob detector. Based on space-scale theory, a multi-scale blob detector was proposed by Lindeberg [88], which could automatically select the appropriate scale for an observation. The scale space can be built using Laplacian of Gaussian (LoG) and difference of Gaussians (DoG) filters. A brief framework for the invariant scale blob detector based on LoG is given by

$$\bigtriangledown^2_{\text{norm}} g = \sigma^2 \cdot \bigtriangledown^2 g(x, y; \sigma), \tag{5.1}$$

where σ is the standard deviation of the Gaussian $g(x, y; \sigma)$, and the scale-space representation $L(x, y; \sigma)$ of the image $f(x, y)$ is defined as

$$L(x, y; \sigma) = \bigtriangledown^2_{\text{norm}} g * f(x, y), \tag{5.2}$$

$$(\hat{x}, \hat{y}; \hat{\sigma}) = \arg[\text{extremum}_{(x,y;\sigma)} L(x, y; \sigma)], \tag{5.3}$$

where $(\hat{x}, \hat{y})$ corresponds to the center vector and $\hat{\sigma}$ to the scale vector of the detected blobs on each scale level. We suppose that one blob center $(\hat{x}_1, \hat{y}_1)$ is stable through the scale space, and a unique maximum over scales is given by

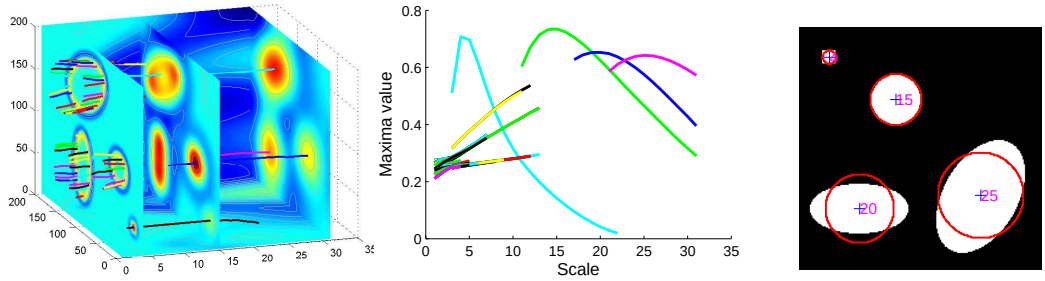
$$\partial_\sigma (L(\hat{x}_1, \hat{y}_1; \sigma)) = 0. \tag{5.4}$$

The evolution of blobs along scales was studied based on the idealized model patterns [88]. In practice, the amount of detected blobs on each scale level is different, and the centers of the same blobs might not be found at the same position on corresponding levels. One common solution is that a blob is detected if a local 3D extreme is present and its absolute value is higher than a threshold [89]. However, same blobs at different scales are not related and can be detected many times. Our strategy is to link the trajectory of the same blobs along scale space and select the scale and center at the unique maxima that best represent each blob. A phantom image as shown in Figure 5.2 (c) that is created for the presentation and the verification of this method. The linked pipelines for each detected blob from the phantom are shown in Figure 5.2 (a). The corresponding local maxima of each pipe through scales are illustrated in Figure 5.2 (b). Figure 5.2 (c) shows the final scale selection by equation 5.4.

In addition, the DoG is a close approximation to the scale normalized LoG, $\bigtriangledown^2_{norm} g$, given by

$$g(x, y; k\sigma) - g(x, y; \sigma) \approx (k - 1) \bigtriangledown^2_{\text{norm}} g, \tag{5.5}$$

where the factor $(k - 1)$ is constant over all scales and has almost no impact on the stability of extrema localization [80]. In this work, DoG is applied for



(a) (b) (c)

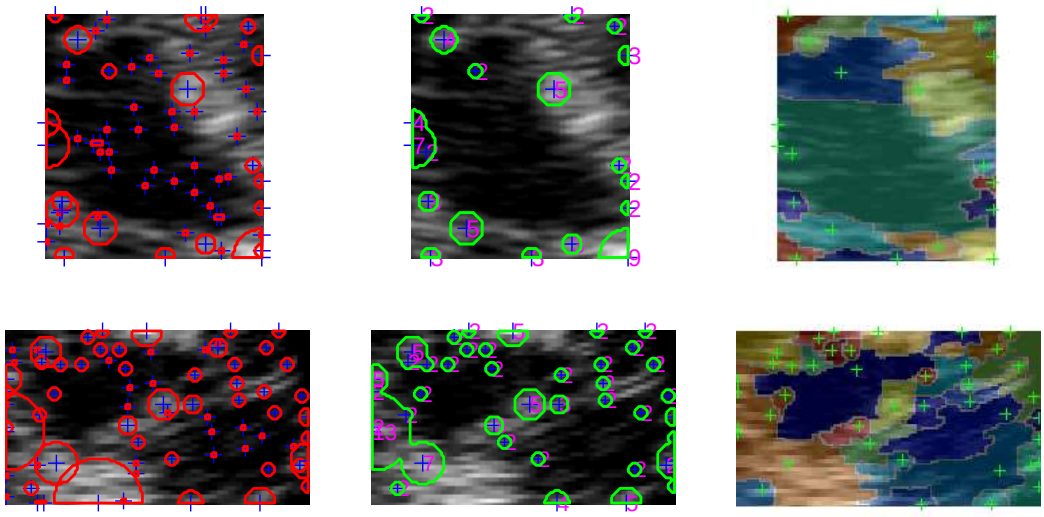
Figure 5.2: Blob scale selection from their trajectories along scale-space representation with LoG. (a) Pipes through three filtered images at scale $\sigma = 4$, 15, and 30 in scale space. (b) Four global maxima at scales 4, 15, 20, and 25 are found from the connected trajectories. (c) Four corresponding blobs are detected and superimposed on the phantom image.

the construction of scale space. Actually, the analysis of scale-space maxima presents severe complications in TCS image, but the possible hyperechogenicity areas are localized by the proposed extrema selection method.

5.2.1 Local Feature Extraction

The mesencephalon is a butterfly-shape-like structure from the transverse view. The TCS image is obtained from the temporal acoustic bone window in a standardized axial mesencephalic imaging plane [27]. Only the half of mesencephalon that is close to the probe is analyzed because of a decreased signal-to-noise ratio with increasing insonation depth. As a result, two TCS images from both sides are acquired per individual. It is better for this study not to include uncertainties that are attributed to the segmentation algorithm. Therefore ROIs are manually segmented by physicians and then analyzed for the estimation of the hyperechogenicity. The hyperechogenicity area is indicated with the blob detection as shown in Figure 5.3. In the next step, a local image descriptor is needed to label the detected blob. The watershed algorithm [90] works on the gradient of an image, which is invariant to the brightness changes of the image. The watershed regions are thus segmented with the input of the detected blobs to estimate the hyperechogenicity in ROI.

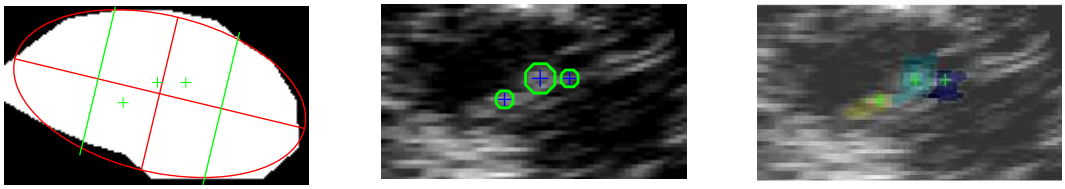
Firstly, the blobs are detected with DoG operators in the ROI using the proposed



(a) (b) (c)

Figure 5.3: The Detected maxima using DoG in (a) 26 neighbors or (b) through pipe in scale space from control images (top row) and PD (bottom row) of Philips SONOS 5500. (c) Watershed segmentation results based on the detected blobs.

extrema-selection method. The detection results of TCS images from Philips SONOS 5500 are shown in Figure 5.3. The same blobs are prevented from being detected many times and the appropriate scales for each blob are indicated around the blob center as shown in Figures 5.3(a) and (b). Secondly, based on the input of the detected blobs, the watershed regions are segmented and labeled by different color as shown in Figure 5.3(c). Then, a selection procedure for the blob and watershed region is implemented with an ellipse mask filtering the false positives as shown in Figure 5.4(a). From the prior knowledge of the anatomic location of SN, this mask is created from the ellipse which is fitted onto the ROI as mentioned in [27]. The values of the ellipse mask are calculated from their distance d to the minor ellipse axis. For $d < f$ (with f the distance between the focus point and the minor axis) the mask value is one. For $d \geq f$ the mask value is zero. Only the blobs that have big scale (for example, $\sigma \geq 3$) are taken into account as shown in Figure 5.4(b). The watershed regions that are entirely within the ROI are considered as interesting areas. As a result, the selected blobs (indicated by green plus signs) and watershed regions are shown in Figure 5.4(c).



(a) (b) (c)

Figure 5.4: The workflow of the interesting area segmentation. From left to right: an ellipse mask is created with the ROI; the blobs are selected by the ellipse mask; the watershed region grows based on the selected blobs. (a) Ellipse mask, two green lines are parallel to the minor ellipse axis and across the two ellipse focuses, respectively. (b) The selected blobs (green sign) and (c) the selected watershed regions which are inside of the ellipse mask.

For the estimation of the hyperechogenicity, nine local features $F1...F9$ were extracted based on the selected blobs and watershed regions in ROI. Entropy was used to measure the randomness of a local region. The parameters shape and scale of a Weibull approximation [91] of the gradient distribution were determined by maximum likelihood estimation [90] and used as local image features. The calculation of entropy and the estimation of Weibull distribution parameters were obtained from the gradient images after Gaussian smoothing. Considering the image scaling, the features F1 and F2 were normalized by the corresponding ROI. The local features are shown as follows:

F1,F3: Area and entropy of all selected watershed regions
F2,F4: Area and entropy of all selected blobs
F5,F6: Weibull parameters (a,b) of all selected watershed regions, blobs
F7: The scale of the biggest detected blob
F8,F9: Entropy of the biggest blob and ROI

The performance of these local features will be presented in the following.

5.2.2 Experimental Results

The experiments were implemented on three data sets which were obtained with Philips SONOS 5500 by different examiners. Dataset 1 includes 42 TCS images

from 23 PD patients and 36 TCS images from 21 healthy controls. Dataset 2 includes 15 PD TCS images from ten PD patients and eight control images from four controls. The last dataset consists of ten PD TCS images from five PD patients and 27 TCS from 14 controls. Totally, this large dataset includes 67 PD images from 38 PD patients and 71 control images from 39 healthy subjects.

The outline of the framework: First, the dataset was classified using the selected feature subsets F(17, 25, 26, 27, 29) from [37] and F(17, 77) from [38]. Secondly, based on the manually segmented ROIs which were marked by the physicians, the suspicious hyperechogenicity areas were localized by the invariant scale blob detection method. Then, the watershed-segmentation algorithm was applied to the gradient image after Gaussian smoothing. At last, local features were extracted based on the selected blobs and the watershed regions. These local features were evaluated by the feature-selection method SFFS. The criterion function of SFFS was the accuracy of the SVM classifier. The training of SVMs was carried out with sequential minimal optimization (SMO) and a linear kernel. The SVM classification results were cross validated with the leave-one-out method.

The feature analysis results are shown in Table 5.1. Based on this dataset, the features found in [37] and [38] achieved 76.81% and 48.55% correct rate, respectively. Five local features F(3,7,8,1,9) were selected with SFFS. Using the selected local featrues, the classification accuracy reached 72.46%, which was better than the Gabor feature and GLCM feature from [38]. To test how the feature sets perform when standard operations such as brightness and contrast normalization are carried out, for each image the intensity values in the ROI were normalized to the range $[0, 255]$. The results in the right column show that the local features are invariant to illumination changes in the simulated images and outperform the other features under such conditions.

Table 5.1: Feature analysis and SVMs cross-validation results on the large dataset.

Dataset	Original data		Normalized data
1, 2, 3	Accuracy	Confusion matrix	Accuracy
$F(17, 25, 26, 27, 29)$ [37]	76.81%	$\begin{pmatrix} 63 & 4 \\ 28 & 43 \end{pmatrix}$	71.01%
$F(17, 77)$ [38]	48.55%	$\begin{pmatrix} 40 & 27 \\ 44 & 27 \end{pmatrix}$	58.70%
Local feature $F(3, 7, 8, 1, 9)$	72.46%	$\begin{pmatrix} 52 & 15 \\ 23 & 48 \end{pmatrix}$	72.46%

In another experiment, a SVM classifier was used to evaluate the performance of the three selected feature subsets when the training was carried out on other datasets than the test. We used Datasets 1 and 3 for Training and Dataset 2 for test. The classification results are listed in Table 5.2. They show that the classifier with the selected local features works better than the others when training and test conditions are different.

Table 5.2: Classification results of the three selected feature subsets.

Training data(Dataset 1,3) Test data (Dataset2)	Accuracy	Confusion matrix
$F(17, 25, 26, 27, 29)$ [37]	65.22%	$\begin{pmatrix} 15 & 0 \\ 8 & 0 \end{pmatrix}$
$F(17, 77)$ [38]	60.87%	$\begin{pmatrix} 14 & 1 \\ 8 & 0 \end{pmatrix}$
Local feature $F(3, 7, 8, 1, 9)$	78.26%	$\begin{pmatrix} 14 & 1 \\ 4 & 4 \end{pmatrix}$

5.3 Shape-Adapted Blob Detection

The two previous works [37, 38] analyzed data from only one ultrasound machine, and the selected features turned out to be sensitive to user settings and the ultrasound machine itself. The work in [32] did not discuss the comparison between the detected area and the SN area labeled manually by physician. Therefore, a shape-adapted interest area detector is implemented to estimate the hyperechogenicity with a large data that is invariant to scale and affine transformations.

Our goal is to extract robust local image descriptors that are invariant to the illumination and contrast changes from the different settings, even invariant to different ultrasound machines. The outline of this algorithm is illustrated with a flow chart in Figure 5.5. The proposed invariant scale blob detection method [32] is applied to localize the echogenicity area in midbrain region. Then the windowed second moment matrix (WSMM) [92] is applied to calculate the parameters of the anisotropic blobs. At last, the interest areas are grouped by a fitted ROI mask with the normalized SN regions that are created according to the prior knowledge of the anatomic location of SN. The extracted local features from the grouping interest areas are used for the hyperechogenicity estimation.

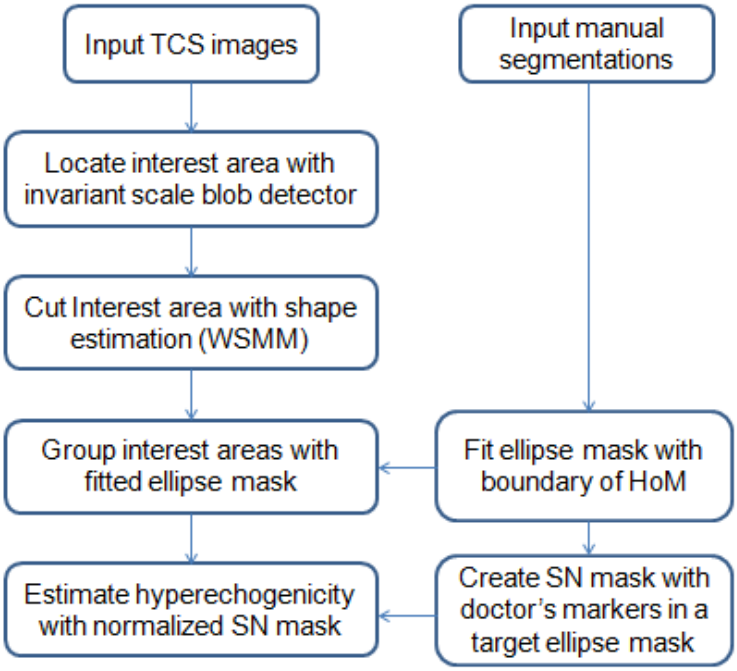


Figure 5.5: The flowchart of image analysis for hyperechogenicity estimation.

The blob detection and watershed segmentation in [32] have been shown to be suited to the TCS images classification. The local features extracted based on that are robust to brightness changes. According to our observations of the data, the hyperechogenic SN area consists of several bright spots in TCS image.

To consider the variety of TCS datasets, an interest area detector must fulfill a number of requirements. First, the detector should be scale invariant so that the bright spots of different size can be localized and extracted. Second, the algorithm should not be sensitive to the different settings of the examiners, such as contrast adjustment and magnification. The invariant scale blob detector used in [32] is stable under the monotonic changes in gray scale. In addition, the blob detector is capable to extract geometric and radiometric attributes, the extracted local features based on that can be used for the classification of TCS images. As a result, in this work, the invariant scale blob detector is applied to localize the echogenicity areas in ROI of TCS image. However, the interest areas could not be surrounded or cut properly with this blob detector. Because the interest areas in ROI are not isotropic based on our observation and it only can be cut with circular boundary edge in [32]. Then the shape information supplied with WSMM is utilized for the edge detection of the anisotropic blobs.

(a) (b)

Figure 5.6: Results of the invariant scale blob detection and the shape adaptation for Gaussian-like blobs. (a) Invariant scale blob detection results. (b) Anisotropic blob detection results with the shape estimation by WSMM.

5.3.1 Shape-Adapted Interest Area Detector

Experiments in Section 5.2 show that the analysis of the scale-space maxima presents severe complications in TCS images, but the possible echogenicity areas could be localized with the blob detector [32] as shown in Figure 5.7 (a). However, the edges of some anisotropic blobs did not mark properly with the blob detector as found from Figure 5.6(a) and Figure 5.7(a). Therefore, the shape information of the elongated blobs must be obtained with other measurement.

Here, a local image descriptor is needed to represent the detected blob. If case the elongated blob has two different characteristic lengths σ_x and σ_y, the unique maximum of $L(x, y; \sigma)$ is obtained at the scale σ that is proportional to $\sqrt{\sigma_x \sigma_y}$. Lindeberg introduced the WSMM to estimate the parameters of a local shape because it fits naturally within the scale-space framework [92].

Let $\bigtriangledown L = (L_x, L_y)^T$ denote the gradient of an image. Given a symmetric and normalized window function w, the windowed second moment matrix $\mu_L(q)$ can be defined as [92]

$$\mu_L(q) = \int_{(x,y)\in\mathcal{R}^2} (\bigtriangledown L)(\bigtriangledown L)^T w(q_x - x, q_y - y) dx dy, \tag{5.6}$$

where $q = (q_x, q_y)$ denotes the pixel at which WSMM is computed. An averaging operator E_q describes the integration with w centered at q:

$$\mu_L(q) = E_q((\bigtriangledown L)(\bigtriangledown L)^T) = E_q \begin{pmatrix} L_x^2 & L_x L_y \\ L_x L_y & L_y^2 \end{pmatrix}. \tag{5.7}$$

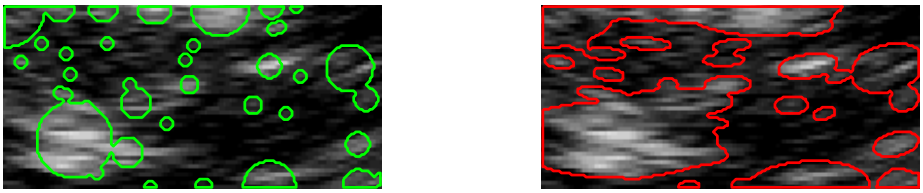
Computation of the WSMM from an image involves two scale problems. The first problem concerns the scale (local scale t) at which to calculate the maxima in the scale-space representation of the image. The second scale problem concerns the size of the region over which to collect the shape information. This scale (integration scale s) controlling the size of the window function (here is selected as a Gaussian function) [92]. Then, the multi-scale WSMM is given by

$$\mu_L(q;t,s) = \int (\bigtriangledown L)(t)(\bigtriangledown L)^T(t) w(q_x - x, q_y - y; s) dxdy. \tag{5.8}$$

In this work, q is the center of the detected blob, t is the selected scale according to the experiment in [92]. In practice, the integration scale s at a given point q is selected as $s = \gamma t$ ($\gamma = 1, \sqrt{2}, 2$). At each pixel, $\mu_L(q;t,s)$ is a 2×2 symmetric positive semi-definite matrix [93]. The shape information of the blob can be easily obtained in terms of the eigenvectors and eigenvalues of the WSMM. Consider a pixel location q, the center of a detected blob with a local scale t, let λ_1 and λ_2 ($\lambda_1 > \lambda_2$) be the eigenvalues of the WSMM, the local neighborhood possesses a dominant orientation (u_2, v_2) [93] which is the eigenvector corresponding to λ_2. When λ_1 and λ_2 are comparable, there is no preferred orientation of this region. Apparently, an ellipse can be defined with the eigenvalues and eigenvectors of the WSMM. The orientation of the major semi-axis of the ellipse is given by (u_2, v_2) corresponding to the smallest eigenvalue λ_2 [94]. Thus, the aspect ratio of the lengths of the semi-axes ρ and the orientations of the detected blobs are defined as [94]:

$$\rho = \sqrt{\frac{\lambda_2}{\lambda_1}}, \quad \theta = arctan(\frac{v_2}{u_2}). \tag{5.9}$$

In the previous work [32], the size of the blob area was set to be proportional to local scale t. For the elongated blob, as similar settings in [94], the width of the elongated blob ω is set to be proportional to t, and the length of the elongated blob is defined by $l = \rho \cdot \omega, (l > \omega)$. As the results detected by the invariant scale blob detector in Figure 5.6 (a) show, the circular boundary calculated with the local scale t was used to mark the detected blobs. Based on the shape estimation, the elongated blob was segmented with an ellipse boundary, the result is shown in Figure 5.6 (b). Obviously, the ellipses created with the shape information are more suitable to mark the anisotropic blobs than the circles. The comparison of the detected results without and with the shape information from a TCS image is shown in Figure 5.7.



(a) (b)

Figure 5.7: Illustration of shape-adapted interest area detector. (a) Detected maxima by the invariant scale blob detector from PD image of Philips SONOS 5500. (b) Interest area detector results with shape estimation.

5.3.2 Interest Area Grouping

As the blobs are detected from midbrain region with the shape-adapted interest area detector in Section 5.3, a selection procedure for the blobs is implemented with an ellipse mask of the ROI and SN areas filtering the false positives. The grouped blobs inside of the SN region and their normalized areas are used for the hyperechogenicity estimation. As the right part of the flowchart in Figure 5.5 shows, first, an ellipse is fitted with the manually marked boundary of ROI for each TCS image (Figure 5.4 (b)). Then, the fitted ellipse is mapped to the target ellipse (Figure 5.4 (a)) with the affine transformation. The results of the affine transform applied to the ROI and manual segmentation are illustrated in Figure 2.8. Using the parameters of the affine transform a mask includes the manual segmentations of ROI, and SN is mapped into the target ellipse as shown in Figures 2.8 (a) and (c). At last, the manually segmented SN regions of TCS images from the group 1 in Dataset 1 are superimposed on the target ellipse mask (Figure 5.8 (a)). This normalized SN mask is applied for the purpose of grouping blobs inside of SN region as shown in Figure 5.8 (b).

Midbrain Mask

In the previous work [32], from the prior knowledge of the anatomic location of midbrain and SN, the mask was created from the ellipse which was fitted onto the ROI as mentioned in [27]. This mask was used for the exclusion of the detected blobs which are outside of the half of mesencephalon region. Furthermore, the window (SN mask) was calculated according to the distance of the pixel location

(a) (b)

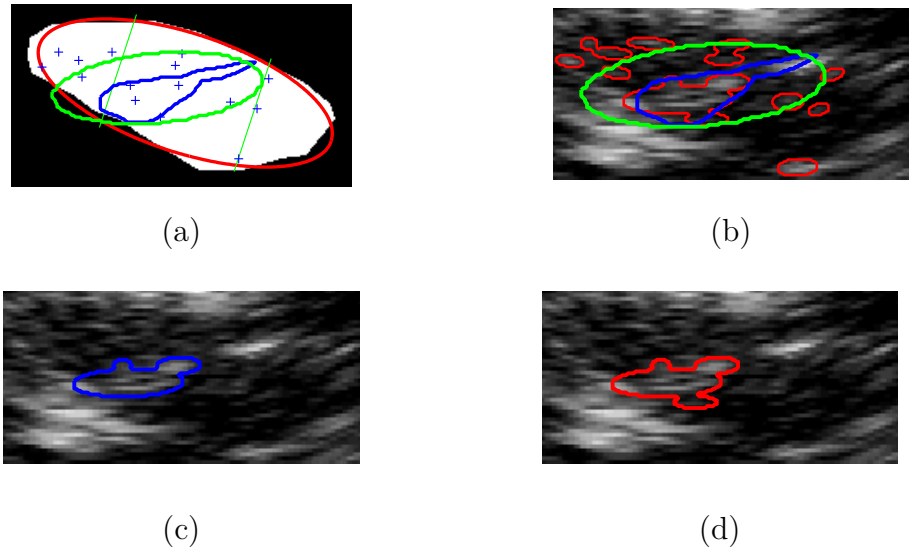
Figure 5.8: The normalization of SN region inside a standard ellipse mask. (a) 23 manually segmented SN regions of PD images are superimposed on the target ellipse mask. (b) The normalized mask (red ellipse) of SN region inside of the half of mesencephalon mask (green ellipse).

d to the minor ellipse axis. For $d < f$ (with f the distance between the focus point and the minor axis) the mask value is one. For $d \geq f$ the mask value is zero. Only the blobs that are within this SN mask were taken into account for the local feature extraction.

Although the window was calculated according to the observation of these TCS images, relative to the anatomic structure of the midbrain, the location of the created SN mask is not accurate. In our study, the strategy is to map each fitted ellipse of each manually segmented boundary of ROI onto a target ellipse with the affine transformation. The centers of the ellipses and the eight control points on the ellipses are applied to calculate the transition matrix of the affine transform. The target ellipse and one fitted ellipse with a ROI are shown in Figure 2.7. The center point and eight control points are indicated in Figures 2.7 (a) and (b). Using the transition matrix, the half of mesencephalon and SN areas of the manual segmentations are transformed onto the target ellipse (Figures 2.8 (a) and (c)). The original image of the half of mesencephalon region and its affine adaptation result can be seen in Figures 2.8 (b) and (d), respectively.

Normalized SN Mask

In order to group the detected blobs, we utilized 23 manually segmented SN regions of PD images from Dataset 1 and mapped them onto the target ellipse mask. Then, 23 affine transformed SN regions were superimposed to create a normalized SN mask. The result is shown in Figure 5.8 (a), and the masks of



(a) (b)

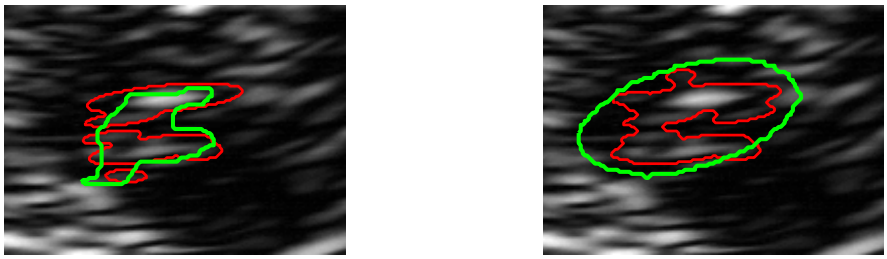
(c) (d)

Figure 5.9: Illustration of the grouping of the interest areas with the normalized SN mask. (a) The detected blobs (blue plus) inside of the manual segmentation of ROI (white region), the manual segmentation of SN (blue curve), and the mapped SN mask (green ellipse) are superimposed in the ROI. (b) The detected interest areas (red curves) inside of ROI are superimposed on the original image. (c) Interest areas inside of the manual segmentation of SN. (d) The interest areas within the mapped SN mask.

the normalized ROI and SN area are illustrated in Figure 5.8 (b). So far, the detected interest areas can be grouped with the normalized SN mask.

First, the transition matrix M of the affine transform from the target ellipse to the fitted ellipse is calculated. Then, the normalized SN mask is transformed onto the ROI of the current TCS image with M. The transformed SN mask and the corresponding manual segmentation of SN are superimposed on the ROI as shown in Figure 5.9 (a). At last, the interest area with the biggest size in the mapped SN mask is selected to estimate the hyperechogenicity as shown in Figure 5.9 (d).

Compared to the manual segmentation of SN, the similar echogenicity are detected with the normalized SN mask from PD TCS images as shown in Figure 5.10. The similar echogenicity can also be found from the TCS images of the healthy controls as illustrated in Figure 5.11.



(a) (b)

Figure 5.10: Illustration of the interest areas grouping on the TCS images of PD from Dataset 2. (a) Interest areas (red curves) inside of the manual segmentations of SN (green curves) or (b) selected by the mapped SN mask (green ellipses).

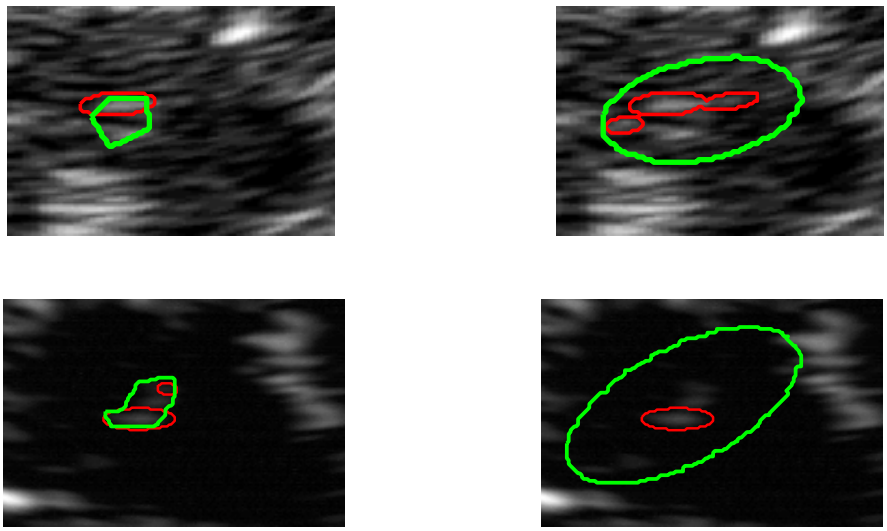
Local Feature Extraction

For the estimation of the hyperechogenicity, five local features $F_l1, ..., F_l5$ were extracted from the midbrain region, SN region, and the selected interest areas within the normalized SN mask. Considering the image scaling, the local features were normalized by the corresponding ROI areas. Entropy is calculated from the gradient images after Gaussian smoothing in order to measure the randomness of a local region. The local features are shown as follows:

Normalized area of manually segmented SN region (F_l1), the selected interest regions (F_l2), and the biggest interest region (F_l3); The entropy of manually segmented ROI (F_l4), the selected interest regions (F_l5).

5.3.3 Experimental Results

The experiments used two datasets which include four groups. The group 1 consists of 21 patients with both PD and Parkin mutation; The group 2 consists of 12 patients with the Parkin mutation; The group 3 includes 16 patients who were diagnosed as PD patients without the Parkin mutation; The group 4 only collects TCS images from the healthy control subjects without the Parkin mutation. Here, a study has been conducted to evaluate whether these local features can be used as a PD indicator and/or a genetic mutation indicator. First, we combined the images of group 1 and group 3 as PD data for the comparison with the healthy controls (group 4). Second, the TCS images in groups 1 and 2 were combined as Parkin mutation data to test whether the



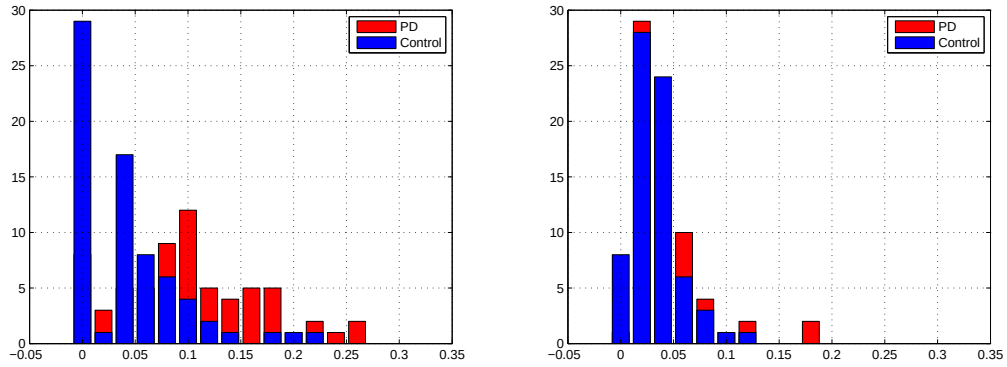
(a) (b)

Figure 5.11: Illustration of the interest areas grouping on the TCS images of the healthy controls from Dataset 1 (upper row) and Dataset 2 (bottom row). (a) Interest areas (red curves) inside of the manual segmentations of SN (green curves) or (b) selected by the mapped SN mask (green ellipses).

local features can separate Parkin mutation carriers from the healthy controls (group 4). In addition, we collected all manually segmented SN areas and then normalized each one with the corresponding ROI area. The histogram of the normalized SN area for each TCS image is shown in Figure 5.12 (a). Similarly, the selected interest areas are normalized with the corresponding ROI area, and the histogram is shown in Figure 5.12 (b).

In the first experiment we compared the diagnosis accuracy between using the manually segmented SN area with the examiners and the interest areas selected by our algorithm. The receiver operating characteristic (ROC) curves were computed as shown in Figure 5.13. If only the SN hyperechogenicity in the manual segmentation of SN was taken into account for the diagnosis of PD, the highest accuracy was 79.82%. The experimental results showed that the performance of the interest areas (77.98%) was not as good as the segmentation of the examiner, the main reason was that the interest areas within the normalized SN mask cannot be as accurate as the marker of the physician.

In the second experiment the normalized SN mask was applied to the PD data,



(a) (b)

Figure 5.12: Illustration of the manually segmented SN areas and the selected interest areas on the TCS images of PD and controls from three datasets. (a) Histogram of the normalized SN area of each TCS image with the crosponding ROI area. (b) Histogram of the normalized areas of the selected interest areas.

the mutation data, and the control data. For each TCS image five local features were extracted with the proposed approach. The performance of the local features was evaluated by the feature-selection method SFFS. The criterion function of SFFS was set as the SVM classifier. The parameter setting of the SVM: a linear kernel, the training was carried out with sequential minimal optimization (SMO), and the cross validation was carried out with the leave-one-out method.

The feature analysis results were shown in Table 5.3 and Table 5.4. Regarding the separation between PD and control data, the selected local feature $F_l(2, 4, 5)$ achieved 84.4% classification accuracy as listed in the first row in Table 5.3. The same feature subset $F_l(2, 4, 5)$ was found to classify the Parkin mutation from the control images, the correct rate is 88.19% as metioned in Table 5.4.

Table 5.3: Feature selection and cross-validation results. Datasets: PD (groups 1 and 3) and control (group 4) data.

Feature subset	Accuracy	Confusion matrix
$F_l(2, 4, 5)$	84.40%	$\begin{pmatrix} 50 & 15 \\ 3 & 41 \end{pmatrix}$
$F_g(28, 12, 27, 92)$	88.99%	$\begin{pmatrix} 59 & 6 \\ 6 & 38 \end{pmatrix}$

In addition, all 101 global features in our previous works [37, 38] were evaluated

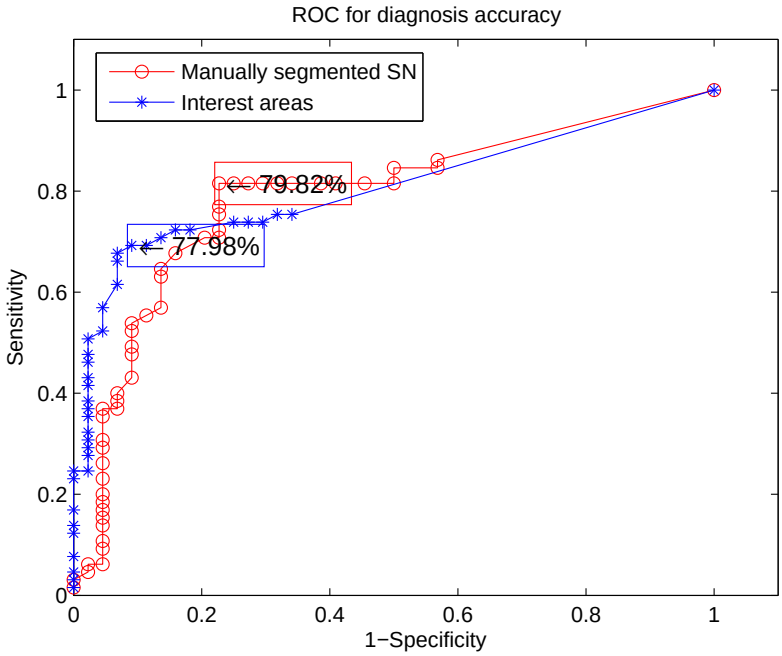


Figure 5.13: The ROC curve of the PD diagnosis accuracy. The red curve with the symbol of circle indicates the manually segmented SN area with the examiners. The blue curve with the symbol of star indicates the selected interest areas.

by SFFS to compare the local features. The selected global feature subset $F_g(28, 12, 27, 92)$ achieved 88.99% to seperate PD images from control images as listed in the second row in Table 5.3. The global features $F_g(27, 28)$ are Gabor features, $F_g(12)$ is the root mean square (RMS) contrast, and $F_g(92)$ is a feature of Gray-Level Co-Occurrence Matrix (GLCM). Similarly, the performance of another global feature subset $F_g(48, 1, 95)$ for the seperation between Parkin mutation and control subjects is 87.13%. Feature $F_g(1)$ is the mean of intensity values of the ROI, the features $F_g(48, 95)$ are one of Gabor features and one of GLCM features, respectively.

Table 5.4: Feature selection and cross-validation results. Datasets: Parkin mutation (groups 1 and 2) and control (group 4) data.

Feature subset	Accuracy	Confusion matrix
$F_l(2, 4, 5)$	88.19%	$\left(\begin{smallmatrix} 48 & 9 \\ 3 & 41 \end{smallmatrix}\right)$
$F_g(48, 1, 95)$	87.13%	$\left(\begin{smallmatrix} 55 & 2 \\ 11 & 33 \end{smallmatrix}\right)$

To test how the feature subsets perform when standard operations such as brightness and contrast normalization are carried out, we applied four normalization methods to each image as mentioned in Section 5.1. The results in Table 5.5 and Table 5.6 show that the local feature $F_l(2)$ is invariant to illumination changes from the image normalization and outperforms the other features under such conditions. The classification rate of the global features decreased to 66.97% from 88.99% with the zero mean and unit variance normalization in Table 5.5. Similarly we also can find the accuracy of the global features dropped from 88.19% to 78.22% in Table 5.6.

Table 5.5: The performance (the percentage of accuracy %) of the feature subset on the simulated data. Datasets: PD and control data.

Simulated data	$F_g(28, 12, 27, 92)$	$F_l(2, 4, 5)$	$F_l(2)$
$[0, 255]$	88.07	78.90	75.23
$\frac{I-\mu}{\sigma}$	66.97	70.64	73.39
Rayleigh	86.24	77.98	75.23
Exponential	86.24	82.57	73.39

Table 5.6: The performance (the percentage of accuracy %) of the feature subset on the simulated data. Datasets: Parkin mutation and control data.

Simulated data	$F_g(48, 1, 95)$	$F_l(2, 4, 5)$	$F_l(2)$
$[0, 255]$	84.16	84.16	78.22
$\frac{I-\mu}{\sigma}$	78.22	78.22	76.24
Rayleigh	80.20	76.24	76.24
Exponential	80.20	88.12	76.24

5.4 Conclusions

We have analyzed the selected features from two previous works and nine new local features based on a large dataset of TCS images. In particular, the local features are invariant to the monotonic changes in gray scale. Almost all possible locations of hyperechogenicity in half of mesencephalon could be indicated by the proposed invariant scale blob detection. Moreover, the watershed segmentation was applied to segment the ROI for PD detection. Of course, the current results depend on the manual segmentation of ROI by physician. Even

though the appearance of mesencephalon can vary considerably across subjects, the prior knowledge of anatomic shape and location of SN can be utilized for the improvement of the selection strategy.

In this work, a large dataset including TCS images of PD patients, Parkin mutation carriers and healthy controls was analyzed with the shape-adapted interest area detector. The elongated blobs on TCS images can be detected with the shape information provided by the WSMM. Moreover, we learned the prior knowledge of anatomic shape and location of SN from the manual segmentations of SN of the physicians and created a normalized SN mask with the affine transform. The possible regions of echogenicity in half of mesencephalon area could be estimated with the normalized SN mask. Based on the detection results, the local features in ROI were extracted and evaluated by the feature selection method SFFS. We have compared the selected feature subsets between the global features in two previous works and the local features based on this large dataset. Regarding the separation between PD and control, Parkin mutation and control, the local features achieved as good performance as the global features. In particular, local features are more stable to the monotonic changes in gray scale and the changes of the histogram shape. Compared to the statistical features, the local feature $F_l(2)$ can provide more information, such as the visual characters (location and area of the echogenicity) on TCS images of PD. However, the detected interest areas depends on the normalized SN mask which varies across subjects. The selection strategy of the interest areas might be improved for the more robust and precise local image descriptors of hyperechogenicity.

Chapter 6

Analysis of TCS Sequence Images

The sequence of ultrasound images known as B-scans are recorded within a fan-shaped region by a movable probe. During the routine examination by TCS, the 2D images than contain the midbrain are chosen by physician for the further diagnosis. In order to reduce the investigator-independence, a 3D image can be reconstructed with a sequence 2D images. The 3D image includes the mesencephalon region can provide more information than a 2D image. In addition, the hyperechogenicity could also be estimated with frames in the sequence that resulted in an automated 2D images selection method for the diagnosis.

6.1 Pre-processing of A Sequence of TCS Image

Reconstruction of volume from a sequence of ultrasound images is an important problem of recent interest [95, 96]. The probe is held by a sonograhper and can be moved and rotated in arbitrary direction. Therefore the probe position and the angular velocity changes slightly during the scanning [95]. The frames need to be registered before the reconstruction, and the similar frames should be combined or merged to avoid the repeated frames affect the quantitive analysis in the next steps.

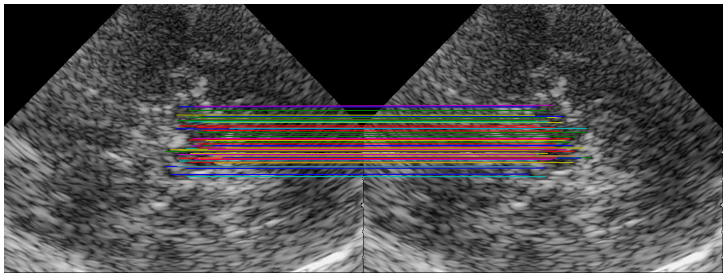
6.1.1 Alignment of the Individual B-scans

The first step is to align the individual B-scans, which are ordered in time domain. We use the SIFT keypoint detector [80], because of its invariance to

image translation and scaling. The keypoint detection and matching of individual B-scans are implemented by VLFeat [97], which is an open source library that implements computer vision algorithms, such as maximally stable extremal regions (MSER) and SIFT. SIFT provides a local descriptor for each keypoint in a pair of TCS images, and then the keypoint descriptors of the pair are matched with the ℓ_2 norm of the difference between them. The matches of the pairs of adjacent TCS images are shown in Figure 6.1(a). Next, the fundamental matrix for the matches are estimated using Random Sample Consensus (RANSAC), and the inliers of the matches that are computed and shown in Figure 6.1(b). At last, the similarity of the pair of images is calculated as the ratio of the inliers to all matches.

(a) 909 tentative matches in a pair of TCS images

(b) 894(98.35%) inlier matches out of 909



(c) The inlier matches inside of masks

Figure 6.1: The SIFT keypoints and the matches between a pair of the adjacent TCS images.

According to the similarity, the adjacent images can be merged using the image mosaicing technique. Image mosaic is to combine two or more B-scans into a new image based on the homograhpy matrix which are computed with RANSAC. Actually, the size of the merged image might be different from the original one. The alignment strategy is to adjust all others images to the first B-scan.

6.1.2 Segmentation of Midbrain in A Sequence

In general, doctors mark the mesencephalon and the SN area on the TCS images for the PD diagnosis. These TCS images are obtained from both sides of the brain as shown in Figure 6.2. Based on the manual segmentation of the physician, the label of mesencephalon in the current B-scan is extracted as a mask (Figure 6.3(a)). Then this mask could be applied to the segmentation of the successive TCS images using the mentioned SIFT features in Section 6.1. The matches inside of the mask are shown in Figure 6.1(c). The matched keypoints in both the current B-scan and the next B-scan (which is adjacent to the current B-scan in time space) are used to calculate the parameters of the affine transformation. Using the similar method as in [80], the mask for the next B-scan is computed from the original mask using the affine parameters.

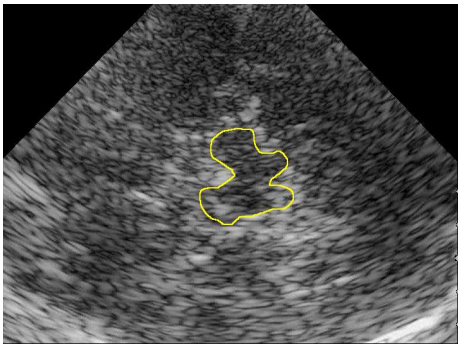


Figure 6.2: TCS image includes midbrain which is indicated by the yellow line.

The affine transformation of a keypoint $[x_1 y_1]^T$ in the current B-scan to the matched keypoint $[u_1 v_1]^T$ can be written as

$$\begin{bmatrix} u_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \tag{6.1}$$

where the affine rotation, scale, and stretch, are represented by parameters m_i, $[t_x t_y]^T$ is the translation parameters. The affine parameters $[m_1 m_2 m_3 m_4 t_x t_y]^T$ are calculated by solving equation (6.1). Therefore, the new mask of the next

B-scan can be transformed from the original mask using the affine parameters. The original mask and the generated mask with the affine transform are shown in Figure 6.3. And then the same processing is implemented to the rest of the B-scan sequence. As a result, the mesencephalon is segmented in all B-scans only based on one manual segmentation from the doctor.

(a) (b)

Figure 6.3: The current mask and the next mask obtained with the affine transform. (a) The current mask. (b) The mapped mask.

6.2 Visualization of Midbrain

Recently, 3D sonography was applied for the diagnosis of PD [98]. The 3D image also shows the hyperechogenicity at SN area, the illustration figure (Abb.4 in [98]) shows the segmentation for the mesencephalon and SN in 3D. In this section, the sequence of B-scans can be reconstructed as a 3D volume data after the pre-processing. An example image is shown in Figure 6.4 that is created based on a sequence images of a healthy control subject.

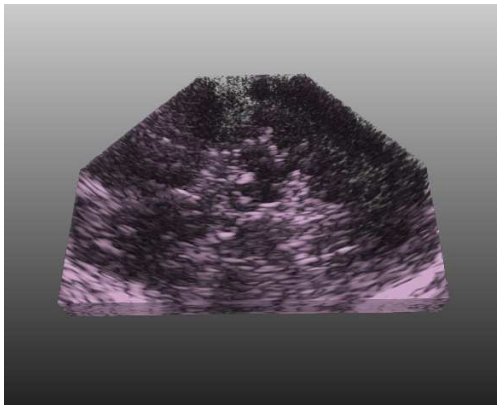
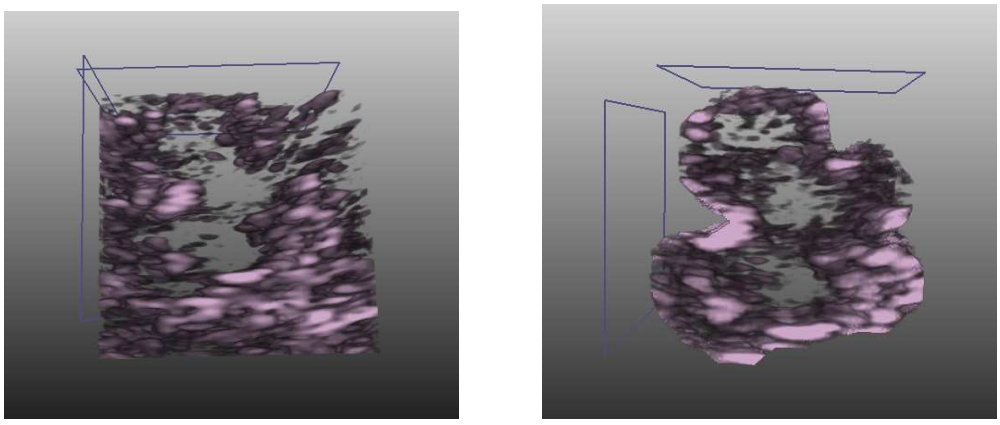


Figure 6.4: The volume rendering result of TCS images after the image fusion and alignment of the frame sequence.

6.2.1 Volume Rendering

After the merging of the similar, adjacent TCS images and the alignment of the sequence, a volume data is reconstructed based on the processed B-scans with the time label. The 3D data is then visualized by the volume rendering, which is supplied with MeVisLab [99]. MeVisLab is a development environment for medical image processing and visualization. The mesencephalon as a butterflyer-shaped structure can be found around the center of Figure 6.4.

Furthermore, in order to obtain a more clear image that mainly includes the midbrain region, the sub-volume of mesencephalon region is clipped from the entire volume as shown in Figure 6.5(a). Obviously the big picture could provide more details in midbrain region. In addition, the volume data build by the segmented frames is visualized in Figure 6.5(b). This 3D image roughly displays the mesencephalic stem, the size of the region will be used for the hyperechogenicity estimation in the next step.



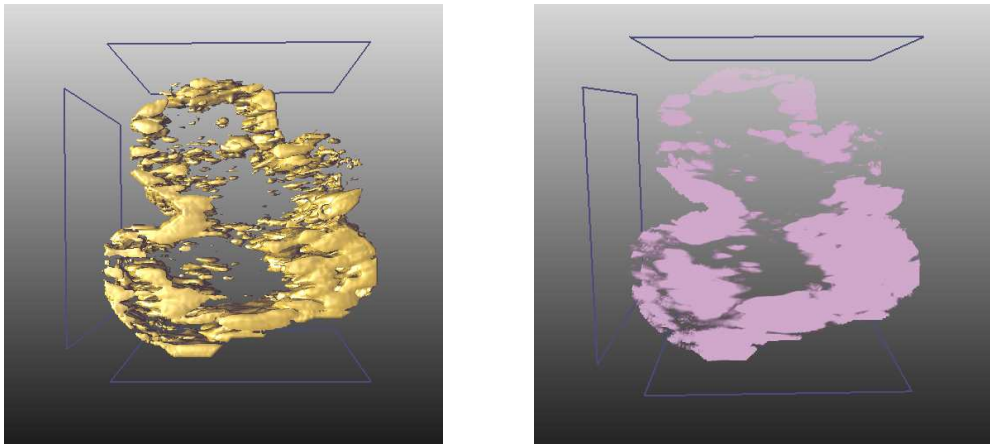
(a) (b)

Figure 6.5: The volume rendering result of (a) the sub-volume of midbrain and (b) the segmented mesencephalon which obtained from one segmented TCS image by the doctor.

6.2.2 Maximum Intensity Projection

According to the standard clinical setting [29], TCS images are acquired at the scanning planes from both sides of the brain. Apparently, 3D images can supply more information than a 2D image of the midbrain. Here, 3D volume images

of the midbrain are created using the mentioned method in section 6.1.2. It is possible to obtain the volume information of mesencephalon and even SN from the 3D image to assist the diagnosis of PD. MeVislab also supplies another two volume rendering methods, the illuminated mode and the maximum intensity projection (MIP). The rendering results are shown in Figure 6.6.



(a) (b)

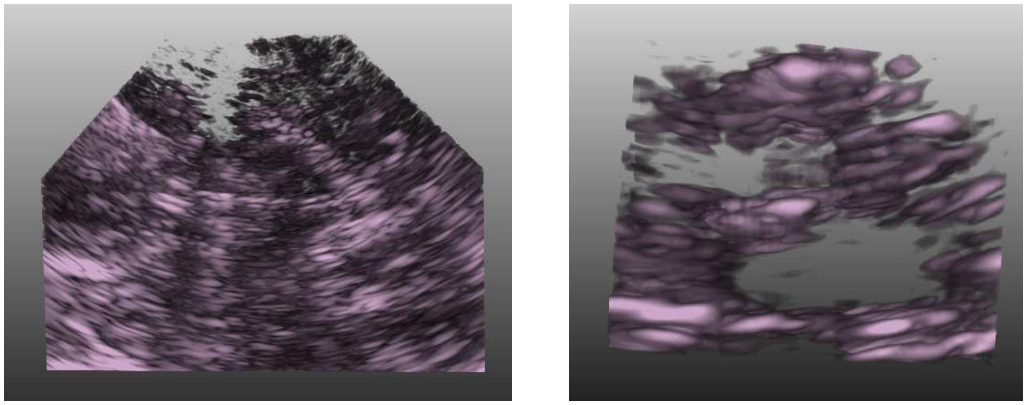
Figure 6.6: The volume visualization using two volume rendering methods. The rendering methods are referred as (a) illuminated and (b) MIP mode from MeVisLab.

Although different parameters can be chosen for the visualization methods in order to obtain better visual image, but the 3D image of mesencephalon in Figures 6.5 and 6.6 show that the intensity values of the SN region are similar to the surrounding area. In other words, the SN region in the upper half of mesencephalon does not show any abnormal in this healthy subject image.

6.3 Analysis on Obtained Sequence

Considering the identification of the scan plane to be investigator independent, consequently, the selection of a 2D TCS image from a sequence of B-scans for the diagnosis also depends on individual physician. Therefore, a sequence analysis method is presented based on the sequence obtained by proposed methods in the previous section. This analysis method aims at selecting 2D images for the diagnosis automatically from the obtained sequence and then the selected images can be used for the estimation of hyperechogenicity. By given a video sequence

from a PD patient, the sequence can be registered by using the SIFT features computed from every two adjacent frames in Section 6.1.1. After merging the similar frames, the resulting sequence is reconstructed into a volume data as illustrated in Figure 6.7 (a). Moreover, the midbrain region on each frame in the obtained sequence can be segmented based on the doctor segmentations as mentioned in Section 6.1.2. The volume data including mesencephalic stem is reconstrcuted based on the frames with the segmented midbrain region and shown in Figure 6.7 (b). As a result, the better diagnosis can be made with the help from 3D visualization of SN region instead of one single 2D TCS image.



(a) (b)

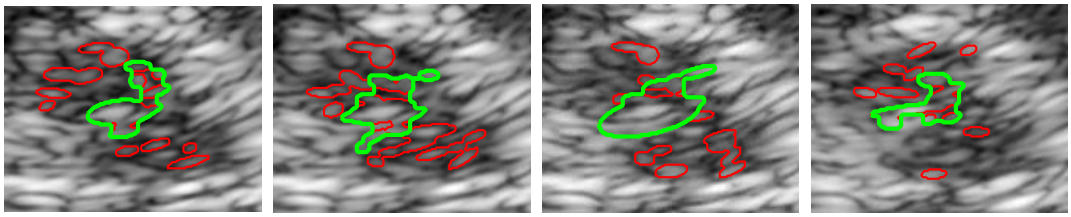
Figure 6.7: The 3D visualization for a PD data after the image fusion and alignment. (a) A data with the whole view of TCS images. (b) A 3D image presets the segmented frames that include upper half of mesencephalon region.

6.3.1 Interest Area Detection

The sequence obtained in last section can be used for the suspect area detection, the hyperechogenicity estimation. In the previous work [32], from the prior knowledge of the anatomic midbrain, the two normalized masks are created from the ellipses as mentioned in Chapter 5 in Section 5.3. One mask is used for the exclusion of the detected blobs which are outside of the half of mesencephalon region. These blobs are detected by the shape-adapted interest area detector that was introduced in Section 5.3.1. Next, the possible areas of echogenicity in this region could be estimated with the normalized SN mask presented in Section 5.3.2. The procedure is implemented on each obtained frame, one by

one, and the two masks are generated by the mapped segmented regions of the midbrain.

The illustration of this processing is given in Figure 6.8. The blobs found by the shape-adapted interest area detector are selected with the mask of half of mesencephalon, that are indicated as red curves and superimposed on the images. The rest blobs are further picked out with the SN mask and the resulting ROIs are displayed as green curves in Figure 6.8. The detected ROIs from all frames in the obtained sequence will be used for selecting proper 2D images for the further diagnosis.



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

Figure 6.8: The illustration of the blobs detection results on the obtained frames that are adjacent in time space. The interest areas (red curves) are detected within the mapped masks of half of mesencephalon. The ROI selected by the SN masks are indicated as green curves. From (a) to (d), the obtained frames in time space.

6.3.2 Identifying of the Scanning Plane

Based on the detection results, the local features in ROI are extracted and then used for the evaluation of each image. The local features show the discriminant power for separating TCS images from PD to the healthy subjects as mentioned in Section 5.3.2. For the estimation of the hyperechogenicity, three local features $F1F2F3$ are calculated based on the regions of the half of mesencephalon and ROI. Considering the image scaling, the feature $F3$ is normalized by $F1$. The local features are shown as follows:

F1: Area of the mapped HoM mask
F2: Area of ROI within the SN mask
F3: ROI normalized by area of HoM mask, the Ratio of F2/F1

The performance of these local features will be presented in the following.

6.4 Experimental Results and Discussion

The data was obtained from the Siemens Sonoline Antares machine, it can output the sequential B-scans as a video file. The video data consists of three healthy control subjects and three PD patients. The sequences are reconstructed in 3D as a volume image. With the help from 3D visualization of SN region, the better diagnosis can be made instead of one single 2D image.

The outline of the framework is: First, every frame including mesencephalon is read out from a video file. Then the image area is extracted from each frame and other information such as text in ultrasound images is erased. Second, the SIFT features are calculated from each image using the software VLFeat. As a result, the keypoints and their corresponding local descriptors are ready for the alignment of the resulting images. In order to reduce the size of the sequence, the similar images are merged according to the similarity between the adjacent frames, which are calculated using RANSAC based on the obtained SIFT features. For example, the current frame is compared to its first neighbor frame in time space, and then compared with its second neighbor frame. The process will be succeeded until the stop condition is fulfilled. In this experiment a threshold of the similarity (96%) is set as the stop condition. The similarity between every two adjacent images of a healthy control is shown in Figure 6.9. The details about the image-merge method are mentioned in Section 6.1.

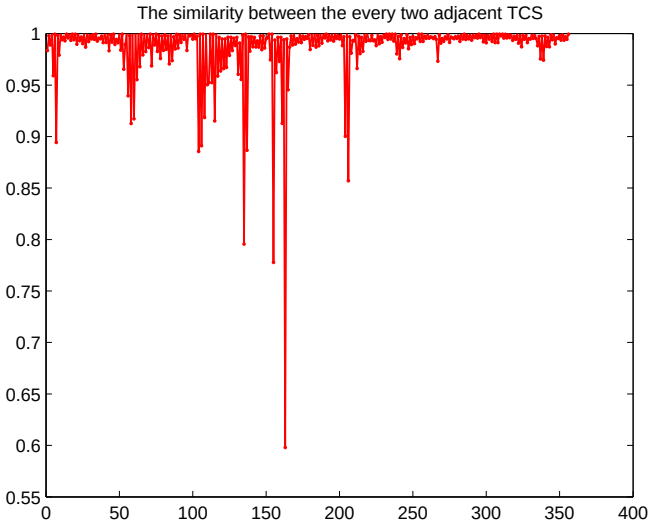
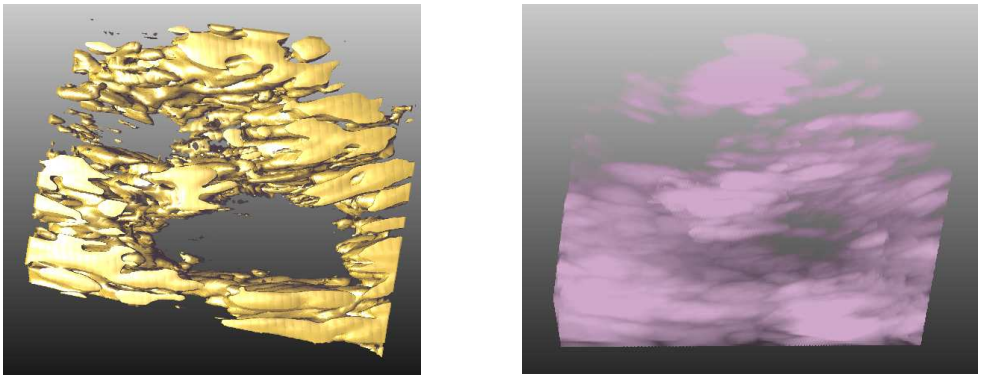


Figure 6.9: The similarity between every two adjacent TCS images in the sequence.

Until this step, the images in the sequence of the healthy control were aligned and the size was reduced from 357 frames to 73 frames. Another example of a PD data, the obtained frames were 69 images that were reduced from 293 original frames. Third, based on the obtained frames and the SIFT features, the manual segmentation of physician on the first frame is mapped by the affine transform from one frame to another. The mapped segmentations are resulted in an atlas for each frame. As a result, the half of mesencephalon region in each frame is extracted as shown in Figure 6.10.



(a) (b)

Figure 6.10: The volume estimation using two volume rendering method, (a) illuminated and (b) MIP mode from MeVisLab.

Next, each obtained frame is analyzed by the shape-adapted interest area detector. The detection results, the blobs, are selected by the normalized SN mask as shown in Figure 6.8. The local feature $F3$ of the ROIs is evaluated and the frames with the large values can be selected for the diagnosis of PD, such as the calculation of the SN area. The feature $F3$ of each obtained frame from the PD data is presented in Figure 6.11.

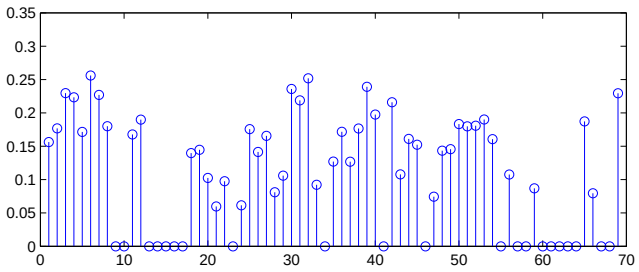
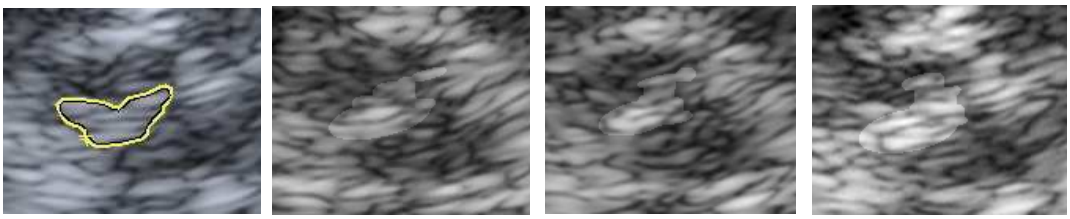


Figure 6.11: The local feature $F3$ for every obtained frame in the sequence.

This procedure could be implemented automatically instead of the manual selection by physician. The comparison between the 2D image selected by doctor and the images selected by the proposed method are shown in Figure 6.12.



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

Figure 6.12: The results of the hyperechogenicity estimation. (a) The B-scan selected by physician, the yellow curve marks the hyperechogenic in SN area. From (b) to (d): The frames (No: 3, 6, and 32) selected by the local feature $F3$. The detected ROIs shown as brighter blobs are overlayed on the original images.

6.5 Conclusions

In this chapter, we apply the local feature SIFT to analyze and process ultrasound image. The keypoints can be matched between a pair of adjacent TCS images. And then the matches are used for the alignment and the image mosaic. Based on one manually segmented TCS image from physician, the segmentation of mesencephalon is implemented sequentially through the sequence of TCS images. As a result, the 3D volume of brain is constructed and the 3D image can be used for the diagnosis of PD. The hyperechogenicity of SN can be evaluated on midbrain region based on the shape-adapted interest area detector. With help of the SN mask, the normalized feature of ROIs is used for the interesting images selection. This automatical selection method could assist physican to recheck whether the scanning plane was proper indentified or the selected images for the diagnosis were suitable.

However, the process of the similarity calculation and the image fusion is time consuming. The accurate measurement of the distance between each frame is still a difficult problem. The hyperechogenicity has been estimated on the limited videos, the accuracy of the quantitative analysis will be improved with a large datasets. With help from physician, the strategy of the images selection for the diagnosis would be more robust.

Chapter 7

Summary and Outlook

Medical imaging plays an essential role in clinical diagnosis. Technical improvements, especially the development of ultrasound device, allowed for the application of transcranial sonographay in the examination of Parkinson's disease. After the revolutionary finding 13 years ago that the ultrsound can penetrate the intact skull with a Doppler system, Becker et al. discovered that the abnormal structure can be detected in the SN region using transcrnial sonography.

This dissertation aims to employ computers to aid neurologist in the diagnosis of PD by using the SN hyperechogenicity as a preclinical marker. With prior knowledge from experts and brain anatomy, this dissertation combines image analysis techniques, machine learning, pattern recognition, and computer vision for the development of computer-aided detection system. Three algorithms have been proposed: a multiple feature extraction, a shape-adapted blob detection, and a sequence analysis approach for PD detection. Especially, the invariant scattering convolution networks is applied on TCS images for the first time and a classification method is proposed that based on the obtained scattering features. In addition, a semi-automatic segmentation tool is applied on the images to reduce the investigator dependence. The results of these algorithm have been given and discussed at the end of the corresponding chapters.

This dissertation firstly continues the work of a previous PhD candidate, Christian Kier, on the feature extraction direction. The features include statistical features, geometrical features, texture features, local features, and scattering features. All these features are extracted from ROI and then used for the classification. The features selected by the feature selection methods provide higher discrimination power on separating PD images from the healthy controls. It is believed that the features combined with the information from the image acquisition stage would be more robust on separating data from different sources.

Valuable information, such as the parameters of user settings, that cannot be obtained from ultrasound machines currently, so that they needs to be treated as unknown parameters.

Actually, the physicians are not fully satisfied with the numbers or percentages provided by the features. Therefore, the shape-adapted blob detector is implemented to find and mark the suspect areas. A series of experiments have demonstrated that the possible hyperechogenic areas can be located based on the TCS images that were obtained from two ultrasound machines, Philips SONOS 5500 and Siemens Sonoline Antares. Moreover, the video data obtained by Siemens Sonoline Antares can be visualized as a 3D image based the sequence analysis method. This tool also could be used by neurologist to recheck the selected 2D images for the quantitative diagnosis. Recently, 3D sonography is applied for the PD diagnosis while 2D imaging techniques are still being used.

Overall, it is believed that one day the neurologist will have a ultrasound device that is provided with a specific probe for TCS, meanwhile a particular designed brain imaging technique is built inside that only focuses on imaging the midbrain region. All PD patients will benefit from the early diagnosis using transcranial sonography.

Appendix

1 The Experiment of Applying DCT on the Scattering Coefficients

Table 1: Percentage of errors (%) of all coefficients and the selected coefficients on the MNIST database. The third column shows the error rates for the selected coefficients with PCA.

Training		DCT Coefficient	DCT4 Coeff	DCT Image	DCT4 Image
300	A	5.23	5.23	12.24	7.60
	B	4.82(105)	4.89(62)	9.20(141)	6.46(50)
	C	2.36(70 d)	2.34(50 d)	2.87(60 d)	3.41(70 d)
1000	A	4.95	4.94	8.43	5.81
	B	4.71(80)	4.80(71)	8.04(160)	5.60(76)
	C	1.86(80 d)	1.87(80 d)	2.12(70 d)	2.62(90 d)
2000	A	4.76	4.69	8.13	5.27
	B	4.37(80)	4.55(78)	7.90(136)	5.03(101)
	C	1.59(80)	1.61(80 d)	1.78(130 d)	2.18(60 d)

2 Part of the Selected Features

Table 2: The part of the selected features in Chapter 3.

Feature	Formula
F(11) Energy	$E = \sum_{i=0}^{L-1} [P(b_i)]^2$
F(12) Entropy	$H = -\sum_{i=0}^{L-1} P(b_i) \log_2(P(b_i))$
	if $P(b_i) = 0$ then $P(b_i) \log_2(P(b_i)) = 0$
F(25) Root Mean Square (RMS) contrast	$RMS = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (I_{ij} - \bar{I})^2}$
	$G_{mn}(x, y) = \sum_{\xi} \sum_{\eta} I(x - \xi, y - \eta) \mathbf{g}^*_{mn}(\xi, \eta)$
F(26) Gabor feautre	$f(1) = \mu_{00} = \frac{\sum_x \sum_y \mid G_{00}(x, y) \mid}{P \times Q}$
F(27) Gabor feautre	$f(2) = \sigma_{00} = \frac{\sqrt{\sum_x \sum_y (\mid G_{00}(x, y) \mid - \mu_{00})^2}}{P \times Q}$
F(29) Gabor feautre	$f(4) = \sigma_{01} = \frac{\sqrt{\sum_x \sum_y (\mid G_{01}(x, y) \mid - \mu_{01})^2}}{P \times Q}$
F(76) Gabor feautre	$f(51) = \mu_{41} = \frac{\sum_x \sum_y \mid G_{41}(x, y) \mid}{P \times Q}$
F(77) Gabor feautre	$f(52) = \sigma_{41} = \frac{\sqrt{\sum_x \sum_y (\mid G_{41}(x, y) \mid - \mu_{41})^2}}{P \times Q}$
F(78) Gabor feautre	$f(53) = \mu_{42} = \frac{\sum_x \sum_y \mid G_{42}(x, y) \mid}{P \times Q}$

Bibliography

[1] Guy Sawle, *Movement Disorders in Clinical Practice*, CRC Press, September 1, 1999, First published 1999, reprinted in 2000.

[2] W. R. Gibb and A. J. Lees, "The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease," *Journal of Neurology, Neurosurgery Psychiatry*, vol. 51, no. 6, pp. 745–752, 1988.

[3] J. Jankovic, "Parkinson's disease: clinical features and diagnosis," *Journal of Neurology, Neurosurg Psychiatry*, vol. 79, pp. 368–376, 2007.

[4] Jose A. Obeso, Maria C. Rodrigruz-Oroz, and B. Benitez-Temino et al., "Functional organization of the basal ganglia: therapeutic implications for Parkinson's disease," *Movement Disorders: official journal of the Movement Disorder Society*, vol. 23 (Suppl 3), pp. 548–559, 2008.

[5] JM. Hagenah, C. Klein, and G. Seidel et al., "Substantia nigra hyperechogenicity correlates with clinical status and number of Parkin mutated alleles," *Journal of Neurology*, vol. 254, pp. 1407–1413, 2007.

[6] C. A. Davie, "A review of Parkinson's disease," *British Medical Bulletin*, vol. 86, no. 1, pp. 109–127, 2008.

[7] "What is Parkinson," http://www.parkinsonsvic.org.au/about-ps/whatps.html, (Accessed: 14 November 2013).

[8] "Genetics home reference," http://www.ghr.nlm.nih.gov/condition/parkinson-disease, (Accessed: 14 November 2013).

[9] S. Behnke, D. Berg, and G. Becker, "Does ultrasound disclose a vulnerability factor for Parkinson's disease?," *Journal of Neurology*, vol. 250 Suppl 1, pp. I24–I2, 2003.

[10] A. H. V. Schapira, "Progress in neuroprotection in Parkinson's disease," *European Journal of Neurology*, vol. 15 (Suppl 1), pp. 5–13, 2008.

[11] J. Spiegel, A. Storch, and Wolfgang H. Jost, "Early diagnosis of Parkinson's disease," *Journal of Neurology*, vol. 253[Suppl 4], pp. IV/2–IV/7, 2006.

[12] David J. Brooks, "Imaging approaches to Parkinson's disease," *Journal of Nuclear Medicine*, vol. 51 (4), pp. 596–609, April 2010.

[13] D. Berg, Jonathan D. Steinberger, C. Warren Olanow, Thomas P. Naidich, and Tarek A. Yousry, "Milestones in magnetic resonance imaging and transcranial sonography of movement disorders," *Movement Disorders: official journal of the Movement Disorder Society*, vol. 26, 2011.

[14] JN. Rutledge, SK. Hilal, AJ. Silver, R. Defendini, and S. Fahn, "Study of movement disorders and brain iron by MR," *AJR American journal of roentgenology*, vol. 149, pp. 365–379, 1987.

[15] P. Mahlknecht, A. Hotter, A. Hussl, R. Esterhammer, M. Schocke, and K. Seppi, "Significance of mri in diagnosis and differential diagnosis of Parkinson's disease," *Neurodegenerative Diseases*, vol. 7, pp. 300–318, 2010.

[16] A. Jon Stoessl, WR. Wayne Martin, Martin J. McKeown, and V. Sossi, "Advances in imaging in Parkinson's disease," *Lancet Neurology*, vol. 10, no. 11, pp. 987 – 1001, 2011.

[17] Annemarie Vlaar, *Transcranial duplex scanning in parkinsonian disorders comparison with SPECT*, PhD thesis, Maastricht University Medical Center, 2009.

[18] M. Eapen, D. H. Zald, and J. C. Gatenby et al., "Using high-resolution MR imaging at 7t to evaluate the anatomy of the midbrain dopaminergic system," *American Journal of Neuroradiology*, vol. 32, pp. 688–694, APR 2011.

[19] D. A. Feinberg and K. Oshio, "GRASE (gradient- and spin-echo) MR imaging: a new fast clinical imaging technique," *Radiology*, vol. 181, no. 2, pp. 597–602, 1991.

[20] W. R. Nitz, "Fast and ultrafast non-echo-planar MR imaging techniques," *European Radiology*, vol. 12, no. 12, pp. 2866–2882, 2002.

[21] M. F. H. Schocke, K. Seppi, R. Esterhammer, C. Kremser, W. Jaschke, W. Poewe, and G. K. Wenning, "Diffusion-weighted MRI differentiates the Parkinson variant of multiple system atrophy from PD," *Journal of Neurology*, vol. 58, no. 4, pp. 575–580, 2002.

[22] G. Becker, J. Seufert, U. Bogdahn, H. Reichmann, and K. Reiners, "Degeneration of substantia nigra in chronic Parkinson's diseasvisualized by transcranial color-coded real-time sonograph," *Journal of Neurology*, vol. 45, pp. 182–184, 1995.

[23] G.Becker, U. Bogdahn, and F. Schlachetzki, *Echoenhancers and Transcranial Color Duplex Sonography*, Berlin: Blackwell Science, 1998, English translation by: Suzyon O'Neal Wandrey, Berlin, reprinted in Germany.

[24] D. Berg and A. Gaenslen, "Place value of transcranial sonography in early diagnosis of Parkinson's disease," *Neurodegenerative Diseases*, vol. 7, pp. 291–299, July 2010.

[25] U. Walter, M. Wittstock, R. Benecke, and D. Dressler, "Substantia nigra echogenicity is normal in non-extrapyramidal cerebral disorders but increased in Parkinson's disease," *Journal of neural transmission*, vol. 109, pp. 191–196, 2002.

[26] J. M. Hagenah, K. Hedrich, B. Becker, P. P. Pramstaller, G. Seidel, and C. Klein, "Distinguishing early-onset PD from dopa-responsive dystonia with transcranial sonography," *Journal of Neurology*, vol. 66, pp. 1951–1952, 2006.

[27] C. Kier, G. Seidel, N. Bregemann, J. Hagenah, C. Klein, T. Aach, and A. Mertins, "Transcranial sonography as early indicator for genetic Parkinson's disease," in *Proceedings of the 4th European Conference of the International Federation for Medical and Biological Engineering*, 2008, pp. 456–459.

[28] D. Berg, J. Godau, and U. Walter, "Transcranial sonography in movement disorders," *Lancet Neurology*, vol. 7, no. 11, pp. 1044–1055, 2008.

[29] D. Berg, S. Behnke, and U. Walter, "Application of transcranial sonography in extrapyramidal disorders: updated recommendations," *European journal of ultrasound*, vol. 27, pp. 12–19, 2006.

[30] G. Becker and D. Berg, "Neuroimaging in basal ganglia disorders: perspectives for transcranial ultrasound," *Movement Disorders: official journal of the Movement Disorder Society*, vol. 16, pp. 23–32, 2001.

[31] M. Okawa, H. Miwa, and T. Kondo et al., "Transcranial sonography of the substantia nigra in japanese patients with Parkinson's disease or atypical parkinsonism: Clinical potential and limitations," *Internal Medicine*, vol. 46(18), pp. 1527–1531, 2007.

[32] L. Chen, J. Hagenah, and A. Mertins, "Feature analysis for Parkinson's disease detection based on transcranial sonography image," in *Proceedings of the 15th International Conference on Medical Image Computing*

and Computer-Assisted Intervention, MICCAI 2012. October, 2012, vol. 7512, pp. 272–279, Springer Berlin Heidelberg.

[33] Kenji. Suzuki, *Machine Learning in Computer-Aided Diagnosis: Medical Imaging Intelligence and Analysis*, IGI Global, 2012, 1-524. Web. 18 Apr. 2013.

[34] Nicholas P. Gruszauskas, K. Drukker, and maryellen L. Giger, *Robustness studies of ultrasound CADx in breast cancer diagnosis*, pp. 1–22, IGI Global, 2012.

[35] Yanhui Guo, *Computer-aided detection of breast cancer using ultrasound images*, PhD thesis, Utah State University, 2010, All Graduate Theses and Dissertations.

[36] Mark A. Haidekker, *Image analysis: A perspective*, pp. 1–22, John Wiley & Sons, Inc., 2010.

[37] L. Chen, G. Seidel, and A. Mertins, "Multiple feature extraction for early Parkinson risk assessment based on transcranial sonography image," in *Proceedings of the 17th IEEE International Conference on Image Processing, ICIP*, sept. 2010, pp. 2277–2280.

[38] L. Chen, J. Hagenah, and A. Mertins, "Texture analysis using Gabor filter based on transcranial sonography image," in *Proceedings of the Bildverarbeitung für die Medizin, BVM*. 2011, pp. 249–253, Springer Berlin Heidelberg.

[39] A. AL-Zubaidi, L. Chen, J. Hagenah, and A. Mertins, "Robust feature for transcranial sonography image classification using rotation-invariant Gabor filter," in *Proceedings of the Bildverarbeitung für die Medizin, BVM*. 2013, pp. 271–276, Springer Berlin Heidelberg.

[40] K. Engel and K. Toennies, "Segmentation of the midbrain in transcranial sonographies using a two-component deformable model," in *Proceedings of the 15th Medical Image Understanding and Analysis*, 2008, pp. 3–7.

[41] S-A. Ahmadi, M. Baust, A. Karamalis, A. Plate, K. Boetzel, T. Klein, and N. Navab, "Midbrain segmentation in transcranial 3d ultrasound for Parkinson diagnosis," in *the 14th International Conference on Medical Image Computing and Computer Assisted Intervention, MICCAI*, Gabor Fichtinger, Anne Martel, and Terry Peters, Eds., vol. 6893 of *Lecture Notes in Computer Science*, pp. 362–369. Springer Berlin Heidelberg, 2011.

[42] A. Sakalauskas, A. Lukoševičius, K. Laučkaitė, S. Rutkauskas, and D. Rastenytė, "Automated segmentation of transcranial sonographic images in the diagnostics of Parkinson disease," *Ultrasonics*, vol. 53, no. 1, pp. 111–121, 2012.

[43] O. Pauly1, SA. Ahmadi, A. Plate, K. Boetzel, and Nassir Navab, "Detection of substantia nigra echogenicities in 3d transcranial ultrasound for early diagnosis of Parkinson disease," in *Proceedings of the 15th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2012, vol. Part III LNCS 7512, pp. 443–450.

[44] Y. LeCun, K. Kavukvuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS*, 2010.

[45] J. Bruna and S. Mallat, "Classification with scattering operators," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2011, pp. 1561–1566.

[46] A. Sakalauskas, A. Lukoševičius, and K. Laučkaitė, "Texture analysis of transcranial sonographic images for Parkinson disease diagnostics," *Ultragarsas (Ultrasound)*, vol. 66, pp. 1392–2114, 2011.

[47] C. Kier, C. Cyrus, G. Seidel, U. G. Hofmanna, and T. Aach, "Segmenting the substantia nigra in ultrasound images for early diagnosis of Parkinson's disease," *International Journal of Computer Assisted Radiology and Surgery*, p. s83, 2007.

[48] K. Engel and K. Toennies, "A quality-of-fit function for evaluating deformable model-based segmentations of anatomical structures in medical data," in *Proceedings of the 16th Conference on Medical Image Understanding and Analysis*, 2009.

[49] K.Engel and Klaus D. Toennies, "Segmentation of the midbrain in transcranial sonographies using a two-component deformable model," *Annals of the British Machine Vision Association and Society for Pattern Recognition, BMVA*, vol. 4, pp. 1–12, Oct 2009.

[50] T. F. Cootes and C. J. Taylor, "Statistical models of appearance for medical image analysis and computer vision," in *Prococeedings of SPIE Medical Imaging*, 2001, pp. 236–248.

[51] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," in *Proceedings of European Conference on Computer Vision*. 1998, vol. 2, pp. 484–498, Springer.

[52] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence,PAMI*, vol. 23, pp. 681–685, 2001.

[53] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[54] J. Xiao, S. Baker, I. Matthews, and T. Kanade, "Real-time combined 2D+3D active appearance models," in *Proceedings of the 2004 IEEE computer society conference on Computer vision and pattern recognition*. 2004, CVPR'04, pp. 535–542, IEEE Computer Society.

[55] K. Zuiderveld Karel, "Graphics gems iv," chapter Contrast limited adaptive histogram equalization. Non-Uniform random variate generation, pp. 474–485. Academic Press Professional, Inc., San Diego, CA, USA, 1994.

[56] C. I. Christodoulou, E. Kyriacou, C. S. Pattichis, and A. Nicolaides, "Multiple feature extraction for content-based image retrieval of carotid plaque ultrasound images," in *Proceedings of the 5th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine*, 2006.

[57] E. Peli, "Contrast in complex images," *Journal of the Optical Society of America*, vol. 7, no. 10, pp. 2032–2040, Oct 1990.

[58] John. C. Russ, *The Image Processing Handbook Sixth Edition*, CRC Press, 2010, NorthCarolina State University.

[59] C. H. Chen, L. F. Pau, and P. S. P. Wang, *Handbook of Pattern Recognition and Computer Vision*, World Scientific Publishing Company, 2 edition, 2005, Second edition.

[60] L. O'Goman, Michael J. Sammon, and M. Seul, *Practical Algorithms for Image Analysis, Second edition*, Cambridge University Press, January 21, 2008, Mixed media product.

[61] Ming-Kuei Hu, "Visual pattern recognition by moments invariants," *IRE Transactions on Information Theory*, vol. 8, pp. 456–459, 1962.

[62] J. R. Parker, *Algorithms for Image Processing and Computer Vision*, Wiley publishing, Inc, 2011, Second edition.

[63] J. Sklansky, "Image segmentation and feature extraction," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 8, no. 4, pp. 237–247, 1978.

[64] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, pp. 610–621, 1973.

[65] Mark A. Haidekker, *Texture analysis*, pp. 236–275, John Wiley & Sons, Inc., 2010.

[66] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, vol. 18, no. 8, August 1996.

[67] I. Fogel and D. Sagi, "Gabor filters as texture discriminator," *Biological Cybernetics*, vol. 61, no. 2, pp. 103–113, June 1989.

[68] Devendran. V, H. Thiagarajan, and A. Wahi, "Svm based hybrid moment features for natural scene categorization," in *Proceedings of the International Conference on Computational Science and Engineering*, 2009, vol. 1, pp. 356–361.

[69] J. R. Parker, *Texture and color*, pp. 177–208, Wiley publishing, Inc, 2011, Second edition.

[70] M. Nixon and Alberto S. Aguado, *Feature Extraction and Image Processing, Second Edition*, Academic Press, 2nd edition, 2008.

[71] J. Han and K-K. Ma, "Rotation-invariant and scale-invariant Gabor features for texture image retrieval," *Image and Vision Computing*, vol. 25, pp. 1474–1481, 2007.

[72] D. Zhang, A. Wong, M. Indrawan, and G. Lu, "Content-based image retrival using Gabor texture features," *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, pp. 13–15, 2000.

[73] E. Fernandez, A. Nieto, F. Martin, and S. Guijarro, "Entropy of Gabor filtering for image quality assessment," in *Proceedings of the 7th International Conference on Image Analysis and Recognition, ICIAR*, 2010, vol. Part 1, pp. 52–61.

[74] E. Vazquez-Fernandez, A. Dacal-Nieto, F. Martin, and S. Torres-Guijarro, "Entropy of gabor filtering for image quality assessment," in *ICIAR'10 Proceedings of the 7th International Conference on Image Analysis and Recognition*, 2010, vol. Part I, pp. 52–61.

[75] N. Cristianini and J. Shawe-Taylor, *Computational Methods of Feature Selection*, Chapman and Hall/CRC, 2007, Edited by Hiroshi Motoda and Huan Liu.

[76] L. Yu, H. Liu, and I. Guyon, "Efficient feature selection via analysis of relevance and redundancy," *Journal of Machine Learning Research*, vol. 5, pp. 1205–1224, 2004.

[77] Avrim L. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artificial intelligence*, vol. 97, pp. 245–271, 1997.

[78] P. Pudil, J. Novovicova, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognition Letters*, vol. 15, pp. 1119–1125, 1994.

[79] S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 1265–1278, 2005.

[80] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.

[81] L. Sifre and S. Mallat, "Combined scattering for rotation invariant texture analysis," in *ESANN 2012 Proceedings of European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, April 2012, pp. 25–27.

[82] S. Mallat, "Group invariant scattering," *Communications in Pure and Applied Mathematics*, 2012, http://arxiv.org/abs/1101.2286.

[83] S. Mallat, "Invariant scattering convolution network," *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, February 2012.

[84] E. Tola, V. Lepetit, and P. Fua, "Daisy: An efficient dense descriptor applied to wide-baseline stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, May 2010.

[85] [online] 2012, "Image scattering toolbox v3," http://www.cmap. polytechnique.fr/scattering, (Accessed: 14 November 2013).

[86] M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars," *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, vol. 31(11), pp. 2032–2047, November 2009.

[87] Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification (2nd Edition)*, Wiley-Interscience, 2000.

[88] T. Lindeberg, "Feature detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, pp. Issue 2, 1998.

[89] L. Ferraz and X. Binefa, "A scale invariant interest point detector for discriminative blob detection," in *IbPRIA 2009 Proceedings of the 4th Iberian Conference on Pattern Recognition and Image Analysis*, 2009, pp. 233–240.

[90] F. Meyer, "Topographic distance and watershed lines," *Signal Processing*, vol. 38, pp. 113–125, 1994.

[91] L. Devroye, *Non-Uniform Random Variate Generation*, New York: Springer-Verlag, 1986.

[92] T. Lindeberg and J. Garding, "Shape-adapted smoothing in estimation of 3-D shape cues from affine distortions of local 2-D brightness structure," in *Proceedings of the 3rd European Conference on Computer Vision, (Stockholm, Swede)*, May 1994, vol. 800, pp. 389–400.

[93] C. Carson, S. Belongie, Hayit Greenspan, and J. Malik, "Blobworld image segmentation using expectation-maximization and its application to image querying," *IEEE Transactions on pattern analysis and machine intelligence, PAMI*, vol. 24, pp. 8, August 2002.

[94] A. Salvatella and M. Vanrell, "Blob detection and grouping for texture desription and other applications," Tech. Rep., 2007.

[95] G. Reis, M. Bertram, R. H. van Lengen, and H. Hagen, "Adaptive volume construction from ultrasound images of a human heart," in *Proceedings of the 6th Joint Eurographic - IEEE TCVG conference on Visualization*, 2003, pp. 321–330.

[96] P. Angelelli, I. Viola, K. Nylund, O.H. Gilja, and H. Hauser, "Guided visualization of ultrasound image sequences," in *Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine*, 2010, pp. 125–132.

[97] [online] 2008 A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms," www.vlfeat.org/, (Accessed: 14 November 2013).

[98] J. Hagenah and G. Seidel, "Distinctive image features from scale-invariant keypoints," *Der Nervenarzt*, vol. 27 August, pp. 1–8–8, 2010.

[99] [online] 2011 MeVis Medical Solutions AG., "MevisLab: medical image processing and visualization," www.mevislab.de, (Accessed: 14 November 2013).