Evaluation of Chronic Wounds by Raman Spectroscopy and Image Processing

A Thesis

Submitted to the Faculty

of

Drexel University

by

Xiang Mao

in partial fulfillment of the

requirements for the degree

of

Doctor of Philosophy

March 2012

Dedications

To my family

Acknowledgements

First of all, I would like to thank Dr. Ahmet Sacan for being my advisor and guided me through the last stage of my Ph.D. program. Dr. Sacan gave me essential directions on my work and patiently helped me. His expert knowledge on data mining and computational languages is admiring. I am grateful to his guidance and help.

I thank Dr. Elisabeth S. Papazoglou for admitting me into her lab and advising me for many years. I also thank her for giving me the trust and the opportunity to work on projects that I am interested in. She had strong passion and remarkable insights; I am deeply regretted that she passed away. She will be remembered by me and many others.

I would like to thank Dr. Andres Kriete, Dr. Hualou Liang, Dr. Kambiz Pourrezaei, and Dr. Somdev Tyagi for serving on my committee and offering valuable suggestions.

Many thanks to Dr. Leonid Zubkov, Dr. Michael Neidrauer, Dr. Chetana Sunkari, and Joshua Samuels for their help on my project. I also thank Dr. Michael Weingarten for the samples and images we acquired from his clinic.

During these years at Drexel, I met friends through various occasions and activities with great pleasure. I am grateful to all my friends for their company.

Finally, I deeply thank my parents and my sister for their love and support.

Table of Contents

List of Tables

List of Figures

Abstract

Evaluation of Chronic Wounds by Raman Spectroscopy and Image Processing
Xiang Mao
Ahmet Sacan, Ph.D.
Elisabeth S. Papazoglou, Ph.D.

Diabetic foot ulcer has become a major healthcare problem as the prevalence of diabetes and the related complications increase globally. Due to the underlying pathological abnormalities in diabetic patients, these ulcers do not heal in a timely and orderly fashion as acute wounds do. Objective and accurate assessment of wound healing status is needed to deliver better wound care to patients.

In this research, we utilize near-infrared Raman spectroscopy to study tissue samples from diabetic foot ulcers on a small cohort of patients. We categorized wounds as healing or non-healing, harvested samples from wound debridement and collected Raman spectra from cryosectioned samples. The average spectrum of samples from healing wounds shows higher intensities at bands associated with collagen and other proteins while the non-healing group shows higher intensities at bands associated with red blood cells. Significant spectral features such as individual band intensities and pairwise intensity ratios were identified by performing unpaired t-tests between these two groups. Supervised classification using a support vector machine (SVM) classifier was conducted to classify the spectra or samples based on the spectral features. The trained SVM classifier is able to predict a spectrum's category with 85.2% accuracy. The prediction of whether a sample is from a healing or non-healing wound can be as accurate as 95.7% when the average spectrum of the sample was fed to the SVM classifier.

Since the quantification of the wound area is a common clinical practice, we also applied image processing techniques to accurately detect the wound boundary in digital images of the

wound. Our method derives from a combination of color based image analysis algorithms, and the method is validated by comparing the performance with manually traced boundaries of wounds in animal models and human wounds of diverse patients. Images were taken by an inexpensive digital camera under variable lighting conditions. Approximately 100 patient images and 50 animal images were analyzed and high overlap was achieved between manual tracings and calculated wound areas by our method. The simplicity of our method combined with its robustness suggests that it can be a valuable tool in clinical wound evaluations.

Chapter 1: Overview and Objectives

1.1 Motivation

Diabetes and diabetic complications have become a serious global health problem. In 2010, it is reported that 285 million people globally have been affected with diabetes (diagnosed and undiagnosed) and this number is expected to reach 439 million in 2030 (George, Cebioglu et al. 2010). Diabetic foot ulcer is one of the major complications in diabetic patients. About 25 per cent of patients will develop foot ulcers in their lifetime after being diagnosed with diabetes.

Due to various etiologic factors, diabetic foot ulcers do not heal normally as acute wounds do. The healing process is prolonged and impaired, usually remaining in the inflammation phase or healing slowly with a high recurrence rate. These foot ulcers also have as high as 50 per cent infection rate and can result in amputation. The diabetic related cost is also a major burden to the healthcare system. In the United States alone, the direct medical costs of diabetic foot ulcers and amputation was $10.9 billion in 2001 (Boulton, Vileikyte et al. 2005).

Constant monitoring of the status of wound healing is needed in order to choose proper treatment or to evaluate the effectiveness of current treatment. In current wound care clinics, preliminary wound evaluation relies mainly on observations of wound bed by professionals. And the wound healing rate is evaluated by wound size reduction, where the wound size is simply measured by width and length. However, the drawbacks are obvious: human observation can be subjective and unquantifiable, while the measuring of wound size by width and length is not accurate. In addition, no actual biochemical information is collected. To address this issue, improvements on current methods should be made and novel methods should be introduced. Moreover, methods based on optical or imaging technologies are preferred due to the possibility

of being made to be non-invasive, non-destructive, and relatively fast as well as quantifiable measurement methods.

1.2 Objectives and Approaches

The global prevalence of diabetes and its increasing trend has caught a great deal of attention from healthcare providers and biomedical engineers. Engineering approaches, such as medical optics, image processing, spectroscopic methods, and multivariate data analysis, have been introduced to study diabetic foot ulcers or other similar chronic wounds.

The overall goal of this research is to improve the evaluation of the healing status of wounds by applying currently available engineering approaches. The improvement can be achieved in any of the following aspects: collecting more direct or accurate information, obtaining more quantitative results, reducing subjectivity and variations in the process, increasing efficiency, minimizing the need for manual operation and reducing the cost.

More specifically, the goal is approached by applying Raman spectroscopy to gather biochemical information from wounds, then using pattern recognition and classification methods to classify wounds, and using image processing methods to process digital images of wounds.

Hence, the specific aims for this research are:

Aim 1: Develop a wound boundary detection method for the digital images of wounds, so as to improve surface wound size measurement by using image processing methods.

Aim 2: Utilize Raman spectroscopy to characterize wound samples collected from wound debridement of diabetic foot ulcers, and investigate spectral differences between healing wounds and non-healing wounds.

Aim 3: Develop a classification method that is able to classify the wounds as healing or non-healing based on the Raman spectral data collected from wound debridement samples.

An introduction to the background of this research is given in Chapter 2, describing current prevalence of diabetes and the burden of diabetic foot ulcers, wound healing model for normal acute wounds and impaired healing of diabetic wounds. In chapter 3, an image processing method developed for detecting wound boundary from digital images are described. Chapter 4 and 5 are devoted to address specific aim 2 and aim 3 respectively. Conclusions and future work are given in Chapter 6.

Chapter 2: Introduction

In this chapter, a general background knowledge that is related to this study is given. In section 2.1, current epidemic situation and trend of diabetes and diabetic complications will be briefly mentioned. The basic knowledge of skin structure and composition is introduced in section 2.2. The commonly accepted wound healing model for normal acute wounds is introduced in Section 2.3, while pathological factors that induce foot ulcers and impairs ulcers from healing are briefly reviewed in Section 2.4. The last section gives a non-exhaustive review for some of the optical and image processing techniques that have been studied as tools for diagnosing or evaluating diabetic foot ulcers.

2.1 Burden of Diabetic Foot Ulcers

Diabetes and diabetic complications have become a serious global health problem. It is reported that 285 million people globally have been affected with diabetes (diagnosed and undiagnosed) in 2010 and this number is expected to reach 439 million in 2030 (George, Cebioglu et al. 2010). In United States, the data revealed by Centers for Disease Control and Prevention (www.cdc.gov) shows that 25.8 million Americans are affected by diabetes in 2010, which accounts for 8.3% of American population. And it is estimated that 79 million Americans in the 20+ age group have pre-diabetes (CDC 2011). This trend is not only seen in developed western countries, but also in many developing countries in Asia with an even higher growth rate (George, Cebioglu et al. 2010).

Diabetes (also called diabetes mellitus) is a disease described as "a metabolic disorder of multiple etiologies characterized by chronic hyperglycaemia with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action, or both" (Alberti and Zimmet 1998). Diabetes can lead to complications such as heart disease, blindness, kidney failure, neuropathy, foot ulcers and amputations. It is the leading cause of kidney failure and non-traumatic lower limb amputation, and a major cause of heart disease. The risk for death among diabetic patients is twice that of people without diabetes (CDC 2011).

Foot ulcer is one of the major complications diabetic patients. Survey has showed that approximately 12-25% of diabetic patients will develop foot ulcers during their lifetime (Leung 2007), and about 2-3% diabetic patients develop a foot ulcer each year (Reiber, Lipsky et al. 1998). Foot ulcers are "cutaneous erosions characterized by a loss of epithelium that extends into or through the dermis to deeper tissues. Ulcers resulted from various etiologic factors and are characterized by an inability to self-repair in a timely and orderly manner". Ulcers are featured of being unable "to self-repair in a timely and orderly manner" (Reiber, Lipsky et al. 1998). The

healing process of diabetic foot ulcer is impaired comparing to the healing process of acute wounds such as surgical or other trauma wounds. The healing time of a diabetic foot ulcer is much longer, averaged 11-14 weeks in contrast to 3 weeks for acute wounds (Boulton, Vileikyte et al. 2005). It is reported that 50% of these ulcers are infected (Fisher, Wolcott et al. 2010). Severe infections can lead to amputations. Among these people that have diabetes, 15% will end up have lower limb amputations in their lifetime (Deshpande 2008). And amputations resulted from diabetic foot ulcer accounts for 84% of non-traumatic amputations (Brem, Sheehan et al. 2004). The average direct cost of wound care of diabetic foot ulcers ranges from $993 to $17 519, and the cost is ranged from $16 488 to $66 215 if the amputation is resulted (1998 equivalent). In United States alone, the direct medical costs on diabetic foot ulcers and amputation is $10.9 billion in 2001 (Boulton, Vileikyte et al. 2005). The occurrence of foot ulcers not only became a healthcare burden, it also strongly affects the quality of life for patients. Studies on health-related quality of life have found significant difference between diabetic patients have and don't have foot ulcers, and between diabetic patients with current foot ulcers and with healed foot ulcers (Reiber, Lipsky et al. 1998).

2.2 Skin Structure

A wound is defined as "a disruption of normal anatomic structure and function" (Lazarus, Cooper et al. 1994). Healing, which means "restoration of anatomic continuity and function", is a complex dynamic biological process involves "soluble mediators, blood cells, extra cellular matrix, and parenchymal cells" (Singer and Clark 1999). Acute wounds are wounds that can repair themselves or can be repaired in an orderly and timely process. In contrast, chronic wounds have "failed to proceed through an orderly and timely process to produce anatomic and functional integrity, or proceeded through the repair process without establishing a sustained anatomic and functional result" (Lazarus, Cooper et al. 1994; Singer and Clark 1999).

Skin is a protective layer for the human body against the external environment. It is also the largest organ in human body. Break of the integrity of skin will expose human tissues to the external environment and will possibly lead to infection, body fluid loss, and electrolyte imbalance (Baum and Arpey 2005). Human skin has a lateral three-layer structure, comprised of epidermis, dermis and subcutaneous tissues (Figure 2.1 (a)). Epidermis is a stratified squamous epithelium, consists mainly of basal and suprabasel keratinocytes, and a small amount of melanocytes, Langerhan cells, and Merkel cells. Epidermis is a multilayer structure on the surface layer of skin. The thickness of epidermis various from 0.06 mm to 0.8 mm depending on the location. From the surface to inside, the sub-layers are cornified layer, translucent layer, granular layer, spinous layer, and basal layer (Benson and Watkinson 2011).

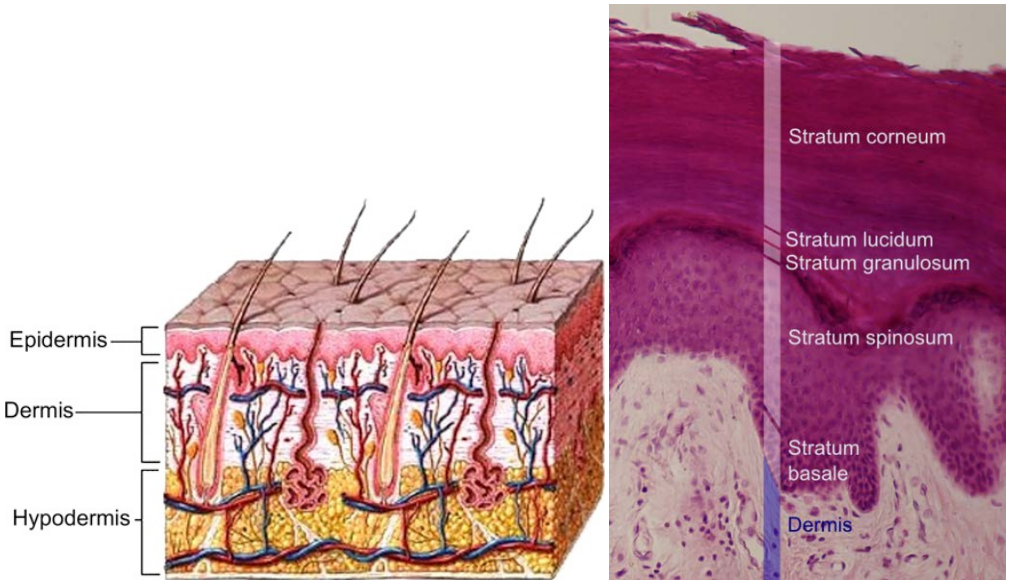

Figure 2.1 (a) The 3 layers of skin: epidermis, dermis, and hypodermis (left). (Cited from (Khavkin and Ellis 2011)). (b) Histologic section of epidermal layers (right). (Cited from (Gantwerker and Hom 2011), originally from Mikael Haggstrom, Uppsala, Sweden; under GNU Free Documentation License. Available at: http://commons.wikimedia.org/wiki/File:Epidermal_layers.png).

Dermis is located underneath the layer of epidermis, about 2-5mm thick, and its main components are collagen fibrils, elastic connective tissues, mucopolysaccharide matrix, and a sparse cell population such as fibroblasts and mast cells. What also contained in dermis layer are vascular network, lymph vessels, other appendages such as hair follicles, eccrine and apocrine glands(Benson and Watkinson 2011). Collagen is the major component of dermis, accounts for 75% of dry weight of dermis. Of all the collagen in dermis, 70% is type I collagen and 15% is type III collagen. However, in arteries, the ratio is reversed (Stadelmann, Digenis et al. 1998; Rook and Burns 2004).

Hypodermis, also called subcutaneous tissue, is comprised mostly of fat cells and interconnecting collagen and elastin fibers.

2.3 Healing Model for Normal Acute Wound

Wound healing is a dynamic and complicated process with many cellular and biomolecular events taking place in a synchronize fashion. Generally, normal cutaneous wounds in adults heal in about 2-3 weeks. In a healed wound, the function of skin cannot be perfectly restored due to wounds heal by scar formation instead of by regeneration (Falanga 2005).

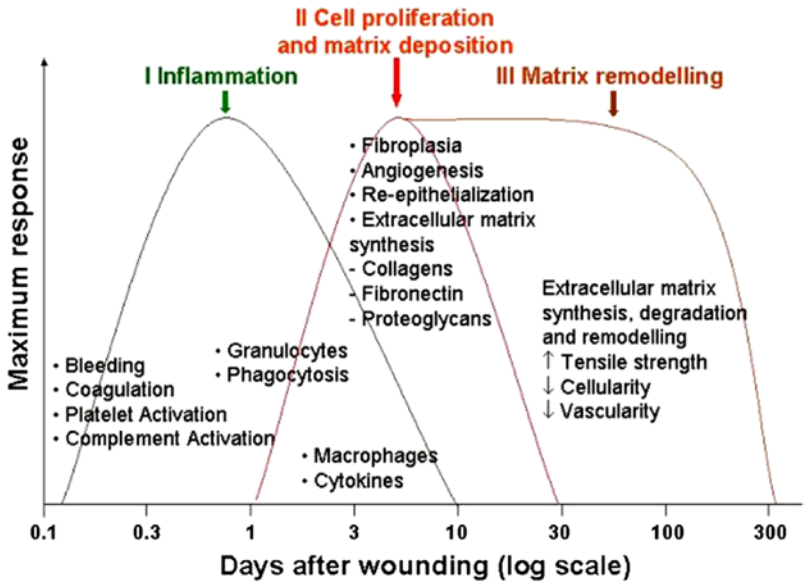

Figure 2.2 Time scale of four overlapped phases of acute wound healing. (Cited from (Gantwerker and Hom 2011), Originally from (Enoch and Price 2004))

In normal acute wounds, the healing process can be generalized as four overlapped phases: hemostasis/coagulation, inflammation, epitheliation and remodeling (Martin 1997; Falanga 2005). Some papers prefer to include hemostasis as part of inflammation phase which happens right after injury (Singer and Clark 1999; Harding, Morris et al. 2002; Baum and Arpey 2005; Gurtner, Werner et al. 2008). The four-phase model will be briefly introduced in this chapter, and a schematic plot that depicts the model is showed in Figure 2.2 (Gantwerker and Hom 2011).

Coagulation happens immediately after the injury in responding to the break of blood vessels. Clots form at injured sites to provide a temporary coverage of the wound and a matrix through which cells can migrate during repairing process. The clot is mainly comprised of platelets and a mesh network that consists mainly of fibrin fibers and other fibers in small amount, such as plasma fibronectin, vitronectin, and thrombospondin. Platelets in the clot can release growth factors and cytokines. Chemotactic cues will also be released to recruit inflammatory cells to the wound site (Martin 1997).

Inflammation also starts right after injury. Neutrophils arrive at the wound site from circulating blood in a few minutes after injury. The functions of neutrophils are not only to clear the initial contaminating bacteria and foreign particles, but also to release pro-inflammatory cytokines that will activate fibroblasts and keratinocytes (Martin 1997). The neutrophil infiltration stops 2 days after the injury if no gross infection happens. Neutrophils are extruded with the eschar or phagocytosed by macrophages (Singer and Clark 1999). Monocytes appear at the wound site2-3 days after injury and become macrophages through metamorphosis (Gurtner, Werner et al. 2008). Macrophages' roles include phagocytosis of bacteria and other cell or matrix debris, and releasing growth factors and cytokines. The inflammation phase can last a few days.

The re-epithelialization phase features with cellular proliferation and cell migration, which take place in 2-10 days after injury (Gurtner, Werner et al. 2008). Formation of ECM proteins, angiogenesis, contraction, and keratinocyte migration are key components in the re-epithelialization phase. As the first event in this phase, epidermis cells, mainly keratinocytes, have to migrate over the interrupted dermis. In order to migrate between collagenous dermis and the fibrin eschar, enzymes such as matrix metalloproteinases (MMPs) and plasmin are needed. MMP-1 (collagenase 1) can degrade type I and III collagen; MMP-9 (gelatinase B) can degrade type IV and VII collagen which are major component in the basement membrane and anchoring fibrils; MMP-10 (stromelysin-2) is able to degrade other non-collagenous matrix components (Martin 1997; Falanga 2005). These MMPs are up-regulated by keratinocytes on wound-edge. Then the angiogenesis process starts, in which new capillaries together with fibroblasts and macrophages forms granulation tissue, providing a new substrate for epidermal cell migration in the remodeling stage in wound repair (Gurtner, Werner et al. 2008). Keratinocytes that are behind the wound-edge become mature and start proliferating, and eventually cover the wound site. It is believed that once a monolayer of keratinocytes forms over the wound site, migration of cells ceases (Martin 1997). Later in this phase, some of fibroblasts differentiate into myofibroblasts

when stimulated by macrophages. Myofibroblasts can help wound contracting, which would speed up wound closure. Fibroblasts and myofibroblasts are both responsible to the production of extracellular matrix (ECM), whose main component is collagen.

Lastly, the remodeling phase takes place 2-3 weeks after injury and might last months or even years. In this phase, all previously describe events dwindle down and eventually cease. Cells such as myofibroblasts, endothelial cells and macrophages either undergo apoptosis or exit from the wound. What remains at wound site are mostly collagen, other ECM proteins, as well as a few dermi cells. This is the process that granulation tissue transforms to scar, and can be dealt as a part of re-epithelialization phase (Singer and Clark 1999). Collagens in ECM are then gradually transformed from type III collagen to type I collagen. This process is mainly controlled by MMPs secreted by macrophages, epidermal cells, endothelial cells as well as fibroblasts. However, healed wound can only attain about 60-70% of strength of uninjured skin (Falanga 2005).

2.4 Diabetic Foot Ulcers and Impaired Wound Healing

Diabetic foot ulcers usually do not heal in a timely and linearly progressed way as acute wounds do. These wounds may remain in inflammation/proliferation phase as the collective result of all various factors. They require much more time to heal, are more prone to infection, are more likely results in amputation caused by severe infection, and have a higher chance of recurrence. As the diabetic population grows, much attention is being directed to understanding the mechanism of diabetic foot ulcers, improving treatment and preventing the occurrence of them.

The causes of ulcers in diabetic patients are multiple. The foremost reason is peripheral neuropathy, which damage their motor, sensory, autonomic fibers. The lack of sensation of pain results in lack of preventive reaction toward pressure or heat, which then leads to trauma, and eventually can develop into ulcers. The other two contributing factors are excessive plantar pressure or foot deformity and trauma (Singh, Armstrong et al. 2005; Leung 2007).

It is reported that more than 100 pathological factors contribute to the impaired healing in diabetic foot ulcers (Brem, Sheehan et al. 2004). Several major factors contributing to the impairment of healing on tissue-molecular level such as neuropathy, vascular problem, immunopathy, and wound infection, are frequently discussed in reviews about diabetic foot ulcers (Calhoun, Overgaard et al. 2002; Jeffcoate and Harding 2003; Falanga 2005; Leung 2007; Velander, Theopold et al. 2008). Meanwhile, extrinsic factors such as callus formation and excessive pressure at wound site are also inneglectable and are in fact related to intrinsic pathological factors.

Neuropathy is considered to be the most prominent factor that leads to ulcer in diabetic patient. Neuropathy, considered mainly due to the abnormal pathway of sorbitol, means damages on sensory, motor, and autonomic fibers (Reiber, Lipsky et al. 1998). Sensory neuropathy can result in trauma without being felt by patients. Motor neuropathy leads to foot deformation and undue pressure on feet. This effect together with ischaemia can lead to subcutaneous tissue necrosis and finally the formation of ulcers with a punched-out appearance (Jeffcoate and Harding 2003). The damage of autonomic nerves can affect microvascular blood flow and change the skin quality.

Vasculopathy is highly associated with non-healing diabetic foot ulcers and amputations. In diabetic patients, the glycation of proteins, such as hemoglobin, albumin, collagen, fibrin, and lipoproteins, are disturbed. This metabolic disturbance affects both macrovascular and microvaslucar activities. In macrovascular, atherosclerosis can happen; while in microvascular, reduction of capillary size, thickening of capillary basement membrane, arteriolar hyalinosis, and impaired endothelial function have been reported (Reiber, Lipsky et al. 1998; Falanga 2005). The overall result from vasculopathy is ischaemia in the tissue, which impedes healing because of insufficient supply of oxygen, nutrients, cells or other biomolecules brought by blood flow.

Infection is a big contributor to impaired healing and amputation in patients with diabetic foot ulcers. The incidence rate of infection in diabetic foot ulcer is also higher than other chronic wounds (Falanga 2005). Immunopathy seems to be closely related to Infections. Although the humoral immunity seems unaffected, these are decreased cellular immunity reported in diabetic patients. The elevated level of glucose concentration is correlated to impaired phagocytosis and the intracellular killing function of the leukocyte (Calhoun, Overgaard et al. 2002).

The metabolic perturbations, due to hyperglyceration, can lead to many abnormalities in cells. It is has been reported that fibroblasts from diabetic foot ulcers showed decrease proliferation upon presence of growth factors. Disturbances to the activity of macrophages are also observed, including decreased release of cytokines and up-regulated MMPs level. There is also evidence of impaired migration of keratnocytes as well as other cell types (Falanga 2005).

As previously mentioned, MMPs play essential roles in the wound healing process. Matrix metalloproteinases (MMPs) is a subgroup of metalloproteinase family. So far 24 MMPs have been identified in mammals (Gill and Parks 2008). All MMPs participate in the catabolism of extra cellular matrix (ECM), and their activities can be inhibited to some extent with the presence of tissue inhibitors of metalloproteinase (TIMPs) (Parks 1999). A notable fact is most of MMPs can degrade or cleave multiple matrix proteins, and the spectrum of proteins that they can catabolize is very similar, while only MMP-1 and MMP-8 appear to be more specific to their substrates (Parks 1999). MMPs are not expressed in uninjured skin. In contrast, MMP levels are present during normal wound repairing process, and it is reported MMP levels are even higher in chronic wounds. A number of studies have detected high level of MMPs in chronic wounds and in chronic wound fluid (Lobmann, Ambrosch et al. 2002; Widgerow 2011). Several studies have suggested that the high value of MMP-9/TIMP-1 ratio is associated with non-healing wounds and it can be used as an indicator of non-healing wounds (Ladwig, Robson et al. 2002; Muller, Belyaev et al. 2007; Liu, Min et al. 2009).

Depending on the underlining major pathological factors, the treatment for diabetic foot ulcer varies (Steed, Attinger et al. 2006; Apelqvist, Bakker et al. 2008). However, the most basic component of diabetic foot ulcer treatment is the wound bed preparation, which includes wound inspection, cleansing, debridement, and blood glucose control. The wound bed preparation, by the definition from “Guidelines for the treatment of diabetic ulcers”, is “the management of the wound to accelerate endogenous healing or facilitate the effectiveness of other therapeutic measures” (Steed, Attinger et al. 2006). Debridement removes necrotic tissue, infected tissue, and calluses from the wound. Proper debridement produces a clean wound surface for better observation, and it also reduces cellular burden of dead cells, bacteria load, and the pressure applied on wound site if calluses are removed. Debridement can be done in several ways, but sharp surgical debridement is preferred (Steed, Attinger et al. 2006). Debridement can be superficial or deep, depending on the severity on infection.

2.5 Current Engineering Approaches

Biomedical engineers are interested in bringing technologies and techniques to help improving diagnosing and treatment. The ultimate goal of this study is to bring in techniques that can improve evaluating wound status.

Patients with diabetic foot ulcer need a comprehensive evaluation of their ulcers. A flowchart by Bremet al. summarized a protocol for treatment of diabetic ulcers in their clinic, which involves several different type of evaluations such as laboratory tests on blood sample, measurement on wound size, ankle-brachial index (ABI, an indicator of blood supply in lower limbs), wound culture (to test if wounds were infected), and other examinations (Brem, Sheehan et al. 2004). In addition to the initial evaluation, the status of the ulcer needs constant evaluation to monitor whether the treatment is effective or not.

The most commonly adopted indicator of wound healing rate is the reduction of wound size. In the guideline for the treatment of diabetic foot ulcers it's recommended that "Patients who fail to show a reduction in ulcer size by 40% or more after four weeks of therapy should be reevaluated and other treatments should be considered" (Steed, Attinger et al. 2006). In Brem's protocol of treatment, they also use the wound size reduction as an indicator of whether the biological therapy should be applied (Brem, Sheehan et al. 2004). The reduction of wound size relative to the initial wound size is considered as wound reduction rate. Some studies have been done on using the reduction rate of wound size as a measure of wound healing speed, or to predict if the wound was going through impaired healing (Jessup 2006; Sheehan 2006). And there is also a whole spectrum of studies related to wound size measurement and analysis. Traditionally, wound size is measured by a ruler and calculated by simply multiplying the length by the width. Wounds can also be traced on transparent films and then the size can be calculated by counting grids. If the wound is not a typical surface wound and depth cannot be overlooked, volume of the wound has to be measured instead of surface area of the wound. In clinics, the depth of a deep wound is measured by finding out how much a cotton swap can stick into the deepest site of the wound. It is not a good estimate of wound volume if only length, width and depth are known. A better measurement of volume is measure how much liquid the wound cavity can hold. However the operation of this measurement can be awkward in some situation. Since taking images of wounds with digital cameras has become a standard procedure of wound archiving in clinics, quite a few studies have been attempted to improve wound size measurement with image analysis (Thawer, Houghton et al. 2002; Wild, Prinz et al. 2008). Belem applied Support Vector Machine, Artifical Neural Network, and Logistic Regression to classify wounds based on the features extracted from wound images (Belem 2004). Duckworth et al. integrated a wound boundary detecting software with a smart phone so that wound boundary can be automatically delineated on screen after wound images being taken (Duckworth, Patel et al. 2007). 3-D cameras that can

measure the volume of deep wounds have also been developed (Plassmann and Jones 1998). Taking digital image of wounds is noncontact and fast. The analyzed results of wound images can be more objective and accurate than human eyes or human empirical judgments given the existence of variations among different individuals. Also, this wound size or color evaluation process is turned more quantifiable by introducing computer aided wound image analysis.

Several other optical and spectroscopic methods can provide more chemical or physiological information have also been introduced to wound healing studies (Neidrauer 2010). For example, laser Doppler perfusion monitoring or imaging (LDPM or LPDI) methods can monitoring blood flow in cutaneous tissue up to 1mm deep by detecting back scattered light from moving particles. Optical coherence tomography (OCT) can image tissue structure up to 1-2 mm in depth by analyzing interference pattern of light signals coming back from the reference arm and the sampling arm. Orthogonal polarization spectral (OPS) imaging can display the distribution of hemoglobin in tissue at a depth about 200 microns. Thermal imaging method has also been applied to study chronic wounds. Diffuse reflectance spectroscopy (DRS) measures the relative concentration of oxy and deoxy- hemoglobin, hence providing information about blood circulation. Deeper tissue spectroscopy can also measure the relative concentration of oxy and deoxy- hemoglobin, but in a relatively deeper depth by applying larger source-detector separations. Diffuse near infrared spectroscopy methodology (DNIRS, can also be called diffuse photon density wave methodology) is able to measure both absorption coefficient and reduced scattering coefficient of the tissue, and is also able to determine the absolute concentrations of oxy and deoxy- hemoglobin.

Despite differences in their methodologies and information they can provide, these optical methods have shared common merits as monitoring tools for biomedical purpose, which are non-invasive, non-destructive, and are relatively fast. However, we should note that these methods cannot provide direct information about chemical composition of wounds. In recent

years, near infrared Raman spectroscopy has caught a lot of attention in biomedical and biological research not only because it can acquire chemical information, but also it is compatible with water and is non-destructive.

Image processing and near infrared Raman spectroscopy are the two main techniques that are applied to study diabetic foot ulcers in this thesis. Chapter 3 will be focusing on image processing of wound images. Chapter 4 and 5 will be focusing on Raman spectroscopy for wound samples and the succeeded data classification. Conclusions and future work will be covered in chapter 6.

Chapter 3: Image Analysis of Chronic Wounds for Determining Surface Area

Progress in wound healing is primarily quantified by the rate of change of the wound's surface area. The most recent guidelines of the Wound Healing Society suggest that a reduction in wound size of less than 40% within 4 weeks necessitates re-evaluation of the treatment. However, accurate measurement of wound size is challenging due to the complexity of a chronic wound, the variable lighting conditions of examination rooms and the time constraints of a busy clinical practice. In this paper we present our methodology to quantify a wound boundary and measure reproducibly the enclosed wound area. The method derives from a combination of color based image analysis algorithms, and our results are validated with wounds in animal models and human wounds of diverse patients. Images were taken by an inexpensive digital camera under variable lighting conditions. Approximately 100 patient images and 50 animal images were analyzed and high overlap was achieved between manual tracings and calculated wound area by our method in both groups. The simplicity of our method combined with its robustness suggests that it can be a valuable tool in clinical wound evaluations. The basic challenge of our method is in deep wounds with very small surface areas where color based detection can lead to erroneous results and which could be overcome by texture based detection methods.

3.1 Introduction

Chronic wounds present an increasing health challenge as the population ages and the incidence of diabetes grows worldwide (Ramsey, Newton et al. 1999; Boulton, Vileikyte et al.). While several wound care treatments may improve healing in a fraction of chronic wound patients, clinicians often need to try several different treatment options before finding one that is effective for a particular wound. For example, there is evidence that synthetic skin grafts and cell therapies improve healing in diabetic neuropathic foot ulcers compared to standard care; however, only roughly 30-50% of patients who received these treatments healed by 12-20 weeks of care (Margolis, Allen-Taylor et al. 2002). The end point of a successful wound treatment is the complete and permanent closure of a wound. In clinical practice the rate of change in wound surface area is the best way to quantify progress in wound healing. During the prolonged healing process of a chronic wound, quantification of the wound healing rate is critical in assessing the efficacy of treatments (Jessup 2006). As detailed in the next section, several ways of measuring surface area are available, ranging from measuring length and width by a ruler to proprietary image analysis algorithms. In this paper we describe a simple and fast algorithm that combines several image analysis tools to quantify wound area and demonstrate the application of our method to complex human wounds.

3.1.1 Current Clinical Paradigm in Wound Measurements

In 2006, the Wound Healing Society issued its *Guidelines for the best care of chronic wounds*, which included the recommendation that the rate of healing in venous, pressure, and diabetic ulcers be evaluated regularly to determine whether treatment is optimal (Robson and Barbul 2006; Steed, Attinger et al. 2006; Whitney, Phillips et al. 2006). In the case of diabetic ulcers, the guidelines specifically recommend that “Patients who fail to show a reduction in ulcer size by 40% or more after four weeks of therapy should be reevaluated and other treatments should be considered” (Steed, Attinger et al. 2006). Studies of both diabetic foot and venous leg

ulcers have shown a positive correlation between the percentage change in wound area after 4 weeks and complete healing after 12 to 24 weeks (Kantor and Margolis 2000; Robson, Hill et al. 2000; Margolis, Gelfand et al. 2003; Sheehan 2006). Accurate measurement of the wound boundary is a critical step in the assessment of wound status.

Two-dimensional wound size has been traditionally assessed by two methods. One is by measuring wound length and width with a ruler. These measurements can be performed rapidly, but may be inaccurate and unreliable, especially for irregularly shaped wounds (Keast, Bowering et al. 2004). A more accurate method to measure wound size relies on tracing the wound perimeter on a transparent film and calculating the wound size either manually or with computer-aided methods. Tracing is generally more representative of the wound area than ruler measurements, but is significantly more time consuming and carries an increased risk of wound contamination and patient discomfort due to contact with the film (Keast, Bowering et al. 2004).

With the ease and low cost of digital photography, common clinical practice includes taking a digital photograph of the wound, tracing the wound boundary with a computer mouse, and calculating the wound size using specialized computer software (Thawer, Houghton et al. 2002). Digital photographic methods can result in improved accuracy over ruler measurements, and reduce the risk of infection or patient discomfort compared to tracing transparent films. However, similar to the tracing of transparent films, the time required to manually trace wound boundaries from computer images may be prohibitive in a busy clinical environment. It is possible however to automate the process of identifying wound boundaries using image processing techniques that improve the accuracy and objectivity of wound size evaluation.

3.1.2 Color Spaces

Colors can be described in various color spaces such as RGB, L*u*v*, L*a*b*, and HSI. Images can be converted from one color space to another using appropriate formulas (Plataniotis and Venetsanopoulos 2000).

The RGB (Red, Green and Blue) color space is the most straightforward color space because it describes color based on the psychophysical fact that the human eye has three kinds of cones that can sense red, green and blue color. It is currently widely used in digital cameras, televisions, and video recorders. In the RGB color space, colors of each wavelength are described with three numbers, indicating intensity values in three primary colors (R, G, and B) (Schanda 2007). However, the sensitivity of the human eye to different colors is highly non-uniform and therefore the RGB color space is not suitable for tasks that need to quantitatively analyze colors in the way that humans can differentiate them. To address this issue, the CIE (*Commission Internationale de l'Éclairage*) introduced two perceptually uniform color spaces in 1976, the L*a*b* color space and the L*u*v* color space, that were designed to permit quantitative analysis of color differences. In the L*a*b* color space, the L* component represents the perceived lightness, indicating change from black to white. The a* and b* channels are chromatic components, indicating green-red and blue-yellow respectively (Plataniotis and Venetsanopoulos 2000; Schanda 2007). The advantage of using the L*a*b* color space is that the perceptual color difference can be represented as the Euclidian distance in the L*a*b* color space between two color vectors, as described in Ref (Schanda 2007). The L*u*v* color space is very similar to the L*a*b* color space, with the L* component representing perceived lightness, while u* and v* represent chromatic components. Color differences can be evaluated in the L*u*v* color space by the Euclidian distance in a similar manner as in the L*a*b* color space. (Plataniotis and Venetsanopoulos 2000; Ohta and Robertson 2005; Schanda 2007).

The HSI (Hue, Saturation and Intensity) color space is also frequently used in image analysis. It is suitable for identifying color in a manner that simulates human vision. Three components are contained in this color model: hue (H), saturation (S), and intensity (I). Hue represents the spectral composition of a color, with a values ranging from 0° to 360° degrees, i.e. both 0° and 360°represent red, 120° represents green and 240° represents blue. Saturation is a ratio ranging from 0 to 1 (0 represent a gray color and 1 represent a pure color). Intensity describes the brightness of a color, ranging from 0 to 1 (0 is a pure black and 1 is pure white). However, quantifying color differences using the HSI color space can be complicated because the values of hue wrap around from 360^{o} to 0^{o}. As a result, the colors represented at the ends of the hue spectrum (360^{o} to 0^{o}) are both very similar shades of red. Furthermore, the values of intensity and saturation are not independent, i.e. changes in the intensity can alter the saturation (Russ 2002).

3.1.3 Existing Computer Algorithms for Wound Tracing

A wound detection algorithm that makes use of the color differences between leg ulcers and the surrounding tissue was proposed in 2001 (Perez, Gonzaga et al. 2001). The RGB images are converted to the HSI color space, and one rectangular region within the wound and three rectangular regions outside of the wound need to be manually selected. The algorithm then determines which channel (R, G, B, H, S, or I) can provide the maximum difference in pixel intensity between the wound and non-wound rectangular regions, and then performs a thresholding operation in this channel. Thresholding is the process of converting an image with multiple pixel intensity levels into a binary image that has only two intensity levels (1 and 0). The authors of the study applied their algorithm to only four wounds, and did not attempt to quantify the accuracy of their method.

An algorithm similar to (Perez, Gonzaga et al. 2001) was used to segment burns from healthy tissue (Acha, Serrano et al. 2005). The authors first converted the images from the RGB

color space to the L*u*v* color space. The Euclidean distance in color space between each pixel and a manually selected burn region was calculated and then processed with a thresholding method. The algorithm was applied to a set of 35 burn images, and quantitative comparison results were given. To our knowledge, this algorithm has not been applied to the problem of chronic wound segmentation.

Kolesnik and Fexa (Kolesnik and Fexa 2004; Kolesnik and Fexa 2006) have used a color image processing algorithm for wound segmentation that classifies pixels as “wound” or “non-wound” using the Support Vector Machine (SVM) classification method. The SVM classification is a machine learning method that requires a training set of images that have been manually segmented into regions of wound and non-wound. A set of rules is established by the SVM classifier based on differences in the color and texture features of training images, and these rules are used to identify regions of wound and non-wound in any subsequent image that is presented to the classifier. The authors used 50 RGB images traced by experts as training data, and then tested their method using 23 new RGB wound images. Their SVM algorithm was able to correctly classify roughly 94% of the pixels as either “wound” or “non-wound”, compared to professional tracings (Kolesnik and Fexa 2006).

In this paper, we aim to develop an algorithm for performing wound segmentation with minimal manual input and high accuracy. Our algorithm uses a combination of both RGB and L*a*b* color spaces, as well as a combination of threshold and pixel-based color comparing segmentation methods. We compared the computer-generated wound segmentation results with manually traced wounds from our animal and human studies and demonstrated that the wound healing curves derived with the computer-generated wound data are highly correlated with the curves from manually traced wound data. This validates the usefulness of our proposed algorithm for following the healing rates of complicated wounds.

3.2 Materials and Methods

3.2.1 Wound Image Acquisition

Wound images were collected from one animal study and one human subject study both carried out at Drexel University from 2006 to 2008 (Weingarten, Papazoglou et al. 2006; Papazoglou, Weingarten et al. 2008; Weingarten, Papazoglou et al. 2008). Approval to conduct the human subject study was obtained from the Institutional Review Board (IRB) at Drexel University College of Medicine that complies with the ethical rules for human experimentation stated in the 1975 Declaration of Helsinki. Approval to conduct the animal study was obtained from the Institutional Animal Care and Use Committee (IACUC) at Drexel University College of Medicine. 56 images from 16 diabetic hairless rats and 99 images of lower limb chronic wounds from 13 diabetic patients were collected. In the animal study, 4 female hairless Sprague Dawley rats, 5-6 weeks old and approximately 150g each, were purchased from Charles River Laboratory (Wilmington, MA) and allowed to acclimate to their surrounding for 4 weeks until they weighed approximately 200g each. A full thickness wound (around 4.6 cm^2) was made using sterile technique in an animal surgical suite. One wound was inflicted on the left side of the dorsal area of each animal, and digital photographs of the wounds were taken on days 3, 5, 8, 10, and 14 after wound surgery. The wounds were nearly closed after day 14. In the human subject study, 13 human subjects with diabetes and chronic wounds were recruited from the Drexel University Wound Healing Center in Philadelphia, PA. All patients were between 18 and 65 years of age, had documented diabetes mellitus for at least 6 months, and had an ankle or foot wound with a minimum surface area of 1 cm^2 that was secondary to the complications of diabetes, including vascular disease and/or neuropathy. All patients received standard wound care, which included weekly or biweekly debridement, treatment with moist wound healing protocols, and offloading when appropriate. In some patients, active wound healing agents such as topical hydrogels,

growth factors, and hyperbaric oxygen were employed. Wound photographs were taken prior to wound debridement on a weekly or biweekly basis.

All wound pictures were taken with commercially available 8-bit digital cameras. A Nikon D100 digital SLR camera with Canfield Twinflash illumination and polarizing filters was used for all animal images and a FujiFilm® FinePix S700 digital camera with cross-polarized filters in front of the flash and the lens was used for all human wounds. The polarized filters reduce light reflection from the wound surface and allow better estimation of wound boundary. Photographs were taken in rooms with the lights shut off and the objects (wounds) were only illuminated by the flash light on the camera. Previous studies (Acha, Serrano et al. 2005) have shown that the flash light on the camera by itself is able to provide adequate illumination for surface wound images. Picture resolution was 3008 x 2000 for the animal images and 3072 x 2304 for the patients' images. All pictures were saved as JPEG files and ruler stickers placed in the imaging plane of the wounds were used as size references.

3.2.2 Manual Wound Measurement

Wound images were visualized on a computer screen where the wound edges were manually traced with a mouse using Adobe Photoshop™. The pixels inside the traced edges were counted, and pixel size was estimated using the ruler placed in the imaging plane of each wound. The surface area of each wound was calculated by taking the product of the pixel count and pixel size. All image processing was performed in MATLAB 7.6.0 (R2008a) using the Image Processing Toolbox™ (V6.1).

3.2.3 Computer Generated Wound Boundary and Wound Image

We have developed an algorithm that can identify the boundary of a wound based on the relative color difference between the wound and the surrounding skin with a minimum amount of user input. Figure 3.1 briefly shows our procedure of processing and analyzing an image, and the

major steps are detailed below. Only parts of steps 1 and 2 involve minimal user intervention; the remaining steps are fully automated.

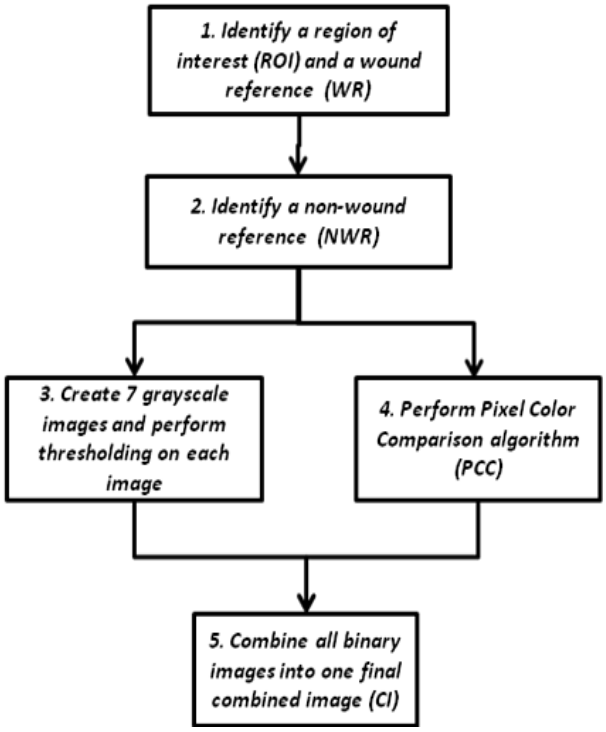

Figure 3.1 Flow Chart of Wound Image Analysis Algorithm.

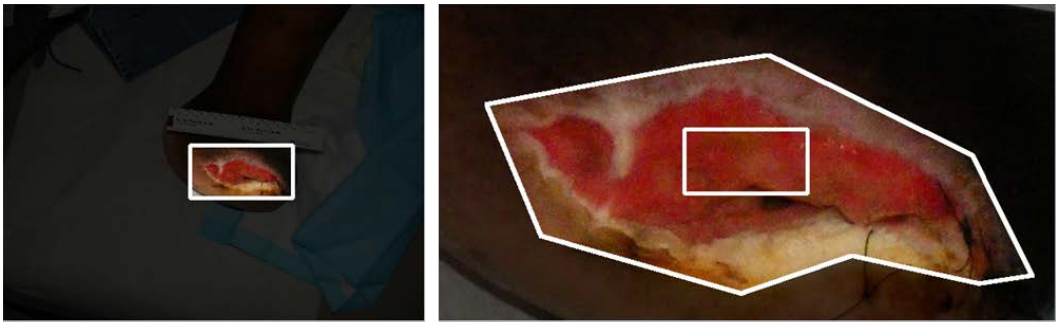

Figure 3.2 (a) Original patient wound image from which the Region of Interest (ROI) is selected (ROI region is shown in its original colors and the remaining part of the image is darkened). (b) A detailed view of the ROI from (a) is shown. The pixels on the perimeter - line of the eight-sided polygon are used as the non-wound reference (NWR). The pixels in the area enclosed by the white rectangle are used as the wound reference (WR).

1. Identification of a region of interest (ROI) and a wound reference (WR): A rectangular region of interest (ROI) is selected from the image so that the wound is centered and occupies most of the cropped image, but there must be intact skin visible on all sides of the wounds. Then a rectangular WR that has approximately 10% of the length and width of the ROI is selected from the center of the wound image.

2. Identification of a non-wound reference (NWR): A polygon is drawn on the intact skin surrounding the wound without including any part of the wound (Figure 3.2). We found that typically 6-8 sided polygons are sufficient, depending on the wound shape. The pixels on the border of the polygon (1 pixel thickness) are used as a NWR region in the subsequent steps.

3. Transformation of 7 grayscale images to binary images and perform thresholding on each image: Three grayscale images are obtained directly from the R, G, and B channels of the ROI. We converted the ROI image from the RGB color space to the L*a*b* color space using standard functions ('makecform' and 'applycform') in MatLab Image Processing Toolbox.

We also obtain one additional grayscale image by using the well-established image analysis method (Acha, Serrano et al. 2005; Schanda 2007) of average color difference ΔE (See Equation (1)). The process we followed was to calculate the color space difference of every pixel in the ROI ($P_{L^*}(i,j)$,$P_{a^*}(i,j)$,$P_{b^*}(i,j)$,) from the average values of pixels in the same color space coordinates in the WR (w_{L^*}, w_{a^*}, and w_{b^*}):

$$\Delta E(i,j) = \sqrt{[(p_{L^*}(i,j) - w_{L^*})^2 + (p_{a^*}(i,j) - w_{a^*})^2 + (p_{b^*}(i,j) - w_{b^*})^2]}, \qquad (1.)$$

where P(i,j) is a pixel at the (i,j) position of the image plane in the ROI. In the end, seven grayscale images are obtained: three from R, G, B, three from L*, a*, b*, and one from the "color difference" ΔE (equation 1). The intensity values in each gray scale image are different. In order to be able to combine information from these different color spaces, we normalized each

grayscale image to a value between 0 and 1 by setting the minimum pixel value inside the polygon to 0 and the maximum pixel value to 1, and by dividing each pixel value inside the polygon by the original difference between maximum and minimum pixel value, according to the equation (2.):

$$p_{normalized}(i,j) = \frac{p(i,j) - p_{min}}{p_{max} - p_{min}}, \quad (2.)$$

where p(I,j) is the original pixel value in the ROI, p_{max} and p_{min} are maximum and minimum pixel values in the ROI, and $p_{normalized}(i, j)$ is the normalized pixel value in the range of [0,1].

Each grayscale image was converted to a binary image using a threshold operation. The threshold level was calculated for each grayscale image using the following equation:

$$\boldsymbol{TL} = \boldsymbol{I}_{WR} - \boldsymbol{c}^{*}(\boldsymbol{I}_{WR} - \boldsymbol{I}_{NWR}), \quad (3.)$$

where ***TL*** is the threshold level, $\boldsymbol{I}_{WR}$ is the average intensity in the wound reference (WR) , $\boldsymbol{I}_{NWR}$ is the average intensity on the polygon lines (NWR) (Figure 3.2) and **c** is the thresholding coefficient. When $\boldsymbol{c = 0.5}$, the threshold is set to the midpoint between the intensity of the wound and the skin references**.** All the intensities in equation (3) are normalized between 0 and 1. The sensitivity of the algorithm to changes in ***c*** is examined in a later section of this paper. After the threshold is determined, we have binary images where grayscale pixels greater than the threshold are assigned a value of 1 and the remaining pixels are assigned a value of 0. In this process the wound region is not set always to black after thresholding (for example, in the a* channel the wound has higher intensity values than the skin). To simplify further processing, some of the binary images were inverted so that pixels similar to the wound reference region are always set to a value of 1, while pixels different from the wound reference region are set to 0.

4. Creation of pixel color comparison (PCC) images: In the PCC algorithm, we separate wound from non-wound pixels on the basis of all three RGB color channels simultaneously. First,

the number of intensity levels in each channel is reduced from 256 to 25 (i.e. the total number of colors is reduced from 3 x 256 to 3 x 25 in RGB) to increase the homogeneity of colors within the ROI. All combinations of red, green and blue color intensities from pixels outside the polygon are considered as the NWR. Next, every pixel inside the polygon is compared to this list of color combinations in the NWR and only those pixels that have color combinations different from the list are classified as wound pixels and assigned a value of 1. Finally, non-wound pixels are assigned a value of 0. The result of the pixel color comparison algorithm serves as an additional (eighth) binary image.

In order to compare the results from images using different methods, we devised a parameter that would allow us to determine how good the fit is between the manually traced area of a wound and the result of our image analysis method. We have adopted a score parameter (**S**), similar to the one used in Ref (Kolesnik and Fexa 2005) and defined by the following equation, as a sensitive measure for this fit.

$$\boldsymbol{S} = \frac{A_{overlapped}}{A_{computed}} \times \frac{A_{overlapped}}{A_{traced}}, \qquad (4.)$$

where $\boldsymbol{A_{computed}}$ is the total number of wound pixels found by our program, $\boldsymbol{A_{traced}}$ is the number of wound pixels defined by the manual trace, and $\boldsymbol{A_{overlapped}}$ is the number of wound pixels that are common to the computer-generated wound area and the manually traced wound area. Therefore, a score of 1 indicates that the computed and traced wound boundaries are identical, while a score of 0 indicates that there are no pixels in common between the computed and traced boundaries.

Scores were calculated for each wound image on all color coordinates with values of c varying between 0 and 1 to evaluate the sensitivity of the wound boundary to different values of c and the effectiveness of each coordinate. Scores from images for all coordinates and for each value of c are grouped together to give an average score and its standard deviation.

5. Combine all binary images into one final combined image (CI): eight binary images are combined to provide a final combined image (CI) of the wound region. This grayscale image is created by calculating a weighted average of all binary images using the following equation:

$$CI = \frac{\sum_{i=1}^{8}(BW_i \times W_i)}{\sum_{i=1}^{8} W}, \qquad (5.)$$

where BW_i represents a binary image in the *ith* channel and W_i represents the weight assigned to the binary image from the *i* th channel. Weights used for generating this combined image are described in the Results Section. The combined intensity image is then converted to a binary image with a threshold of 0.5. This step yields the final binary image of wound area.

3.3 Results

3.3.1 Color Channels and Threshold

The effectiveness of each channel in finding the wound area is evaluated by the score **S**. The arithmetic mean of scores in specified channels from all animal and patient wound images are shown in Figure 3.3, with value of **c** = 0.5. For animal images, all color channels with the exception of the Red channel show high scores **S** (above 0.90). The reason for the low score in the Red channel is due to the fact that in hairless rats both animal skin and wound tissue have small difference in the Red channel. For these simple wound images, it is plausible and convenient to use only the Green or Blue channel to perform wound area analysis.

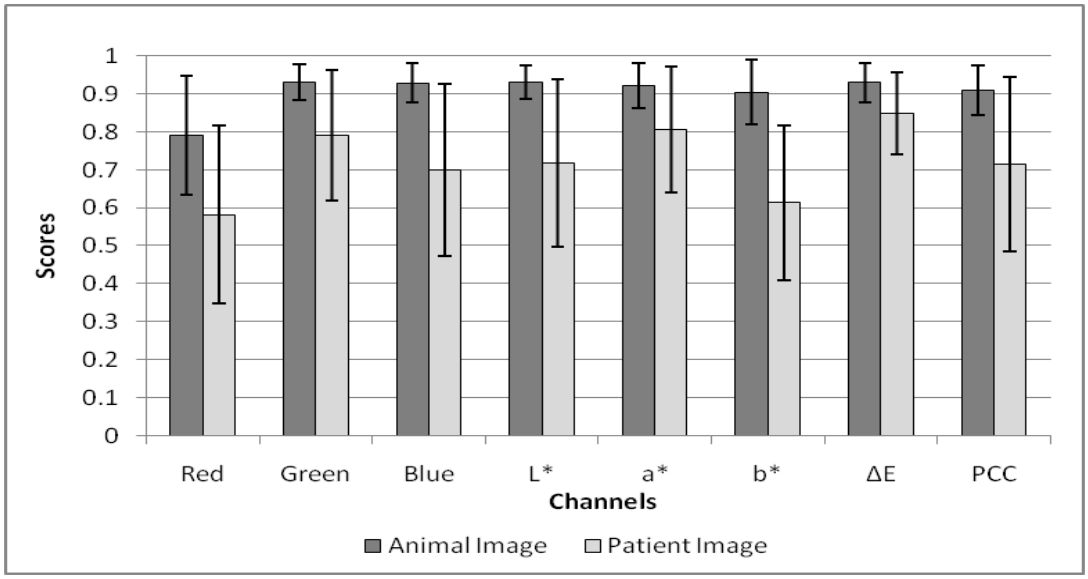

Figure 3.3 Average scores from all animal (dark gray) and patient (light gray) wound images in each channel with a threshold coefficient c = 0.5 Number of animal images N_a = 56, number of patient images N_p = 99.

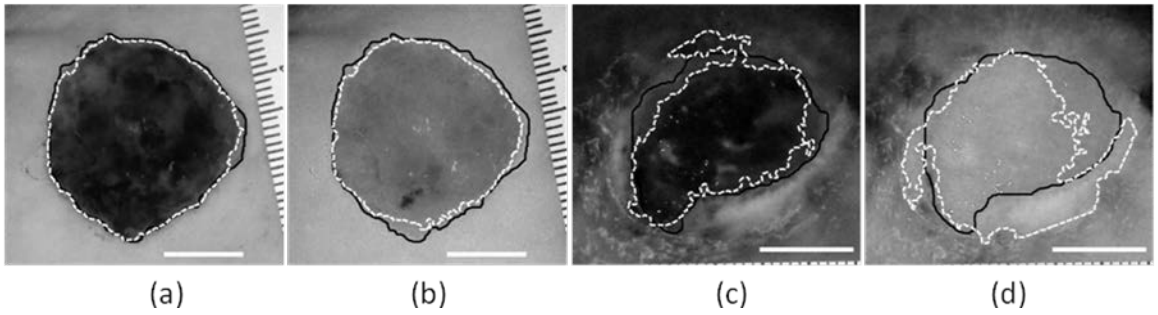

Figure 3.4 (a) Animal image in green channel, S=0.95. (b) Animal image in red channel, S= 0.90. (c) Patient image in green channel S=0.74. (d) Patient image in red channel S=0.53. Black solid lines represent manually traced wound boundaries. White dashed lines represent computer generated wound boundaries when c = 0.5. The white scale bar in each image represents 1 cm.

On the other hand, for patient images, the average color difference Δ*E* channel exhibits the highest score value and the lowest standard deviation (**S** = 0.85±0.11). Channel a* and the Green channel also exhibit high scores and low errors (**S** = 0.81±0.17 for the a* channel and **S** = 0.79±0.17 for the Green channel). The Red channel has the lowest score (**S** = 0.58±0.24), and the b* channel has also low score (0.61±0.20). All other channels have score values between 0.70 and 0.79, with a standard deviation around 0.2. When considering how to combine the results from all channels to provide a final wound image, it is obvious that the information from the Δ*E*,

a* and Green channels is more reliable compared to the Red and b* channels. Examples of wound boundaries obtained from red and green channels for animal and patient wounds are shown in Figure 3.4. Analysis of images from all color coordinates allows us to create a combined image composed of different contributions from the individual coordinates and assess which channels can be used to maximize the score of the final image.

It is important to consider contributions from all color coordinates because for different wounds under variable lighting conditions significant information is contained in these coordinates. Therefore, the eight binary images are combined to provide a final combined image (CI) of the wound region. This gray scale image is created as described in Materials and Methods using equation (5). The weights selected for the contributions of the individual binary images are R =0, G=2, B=1, L*=1, a*=2, b*=0, ΔE =2, and PCC = 1, based on the results shown in Figure 3.3.

The sensitivity of our calculated wound boundaries from the combined images to the threshold coefficient ***c*** is shown for a typical animal image and a typical patient image in Figure 3.6, and the sensitivity of the average scores for all animal and patient images is shown in Figure 3.5. In patient images, the highest score and lowest standard deviation (0.86±0.10) is achieved with a coefficient of 0.5. The scores with a coefficient of ***c*** =0.4 or 0.6 are also very close to the score achieved with a 0.5 value. Our data demonstrate the robustness of the method because the scores remain higher than 0.80 for any value of c between 0.2 and 0.8. For values of c <0.2 and c> 0.8 the score changed approximately by 10%. A similar trend is seen in animal images for all individual coordinates (Figure 3.5).

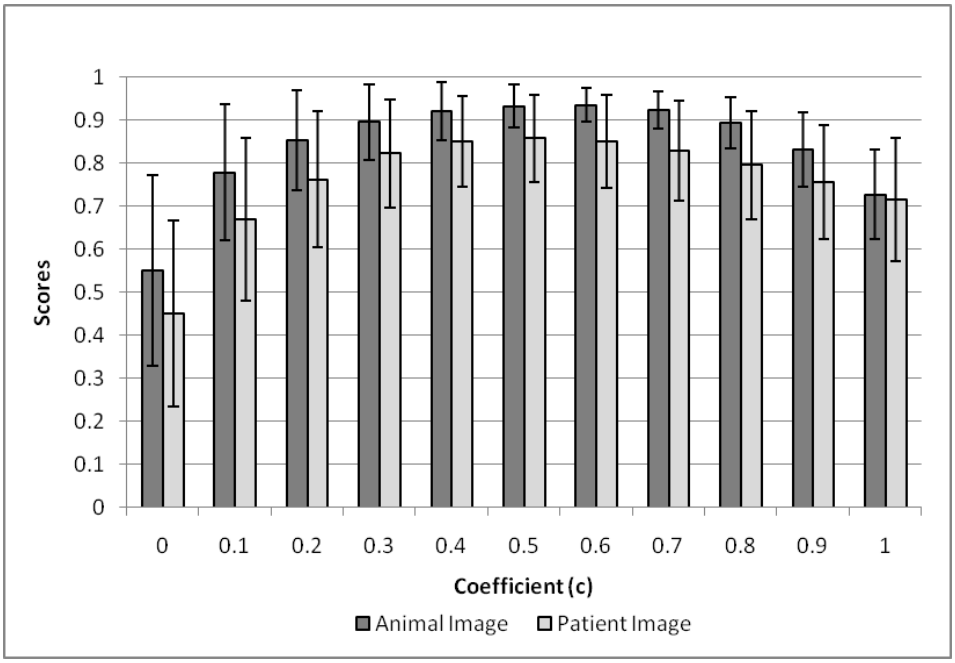

Figure 3.5 Average scores for the combined images: animal results shown in dark gray and patient results in light gray. Number of animal images N_a = 56, number of patient images N_p = 99.

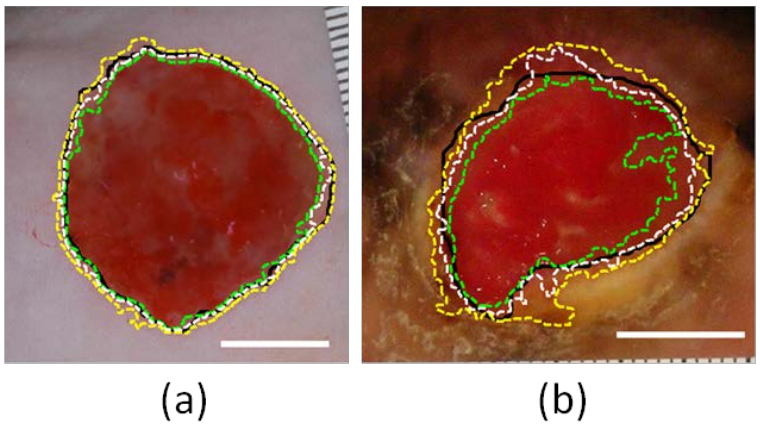

Figure 3.6 Sensitivity of wound area to values of c: Comparison of calculated wound boundaries (combined image) obtained using different values of c with manually traced wound boundaries (Black solid lines). Green dashed, white dashed and yellow dashed lines are computer generated wound boundaries with c = 0.3, 0.5, and 0.8, respectively. (a) Typical animal wound image. The areas enclosed by the calculated wound boundaries are: 4.25 cm^2, 4.55 cm^2, and 4.97 cm^2 for c = 0.3, 0.5, and 0.8, respectively. The area of the manually traced wound boundary is 4.62 cm^2. (b) Typical human wound image. The areas enclosed by the calculated wound boundaries are: 8.43 cm^2, 8.79 cm^2, and 9.30 cm^2 for c = 0.3, 0.5, and 0.8, respectively. The area of the manually traced wound boundary is 8.99 cm^2. The white scale bar in each image represents 1cm.

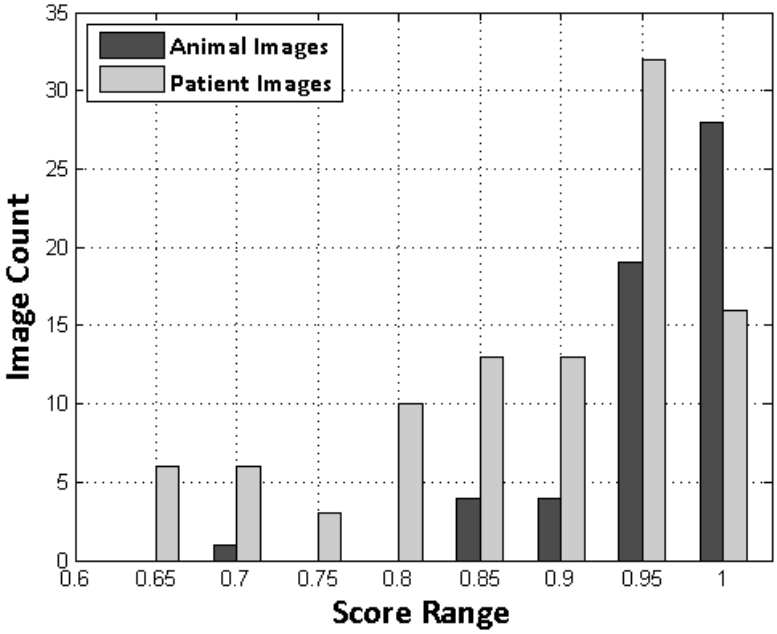

Figure 3.7 Histogram of scores from all animal (dark gray) and patient (light gray) images. Number of animal images N_a = 56, number of patient images N_p = 99.

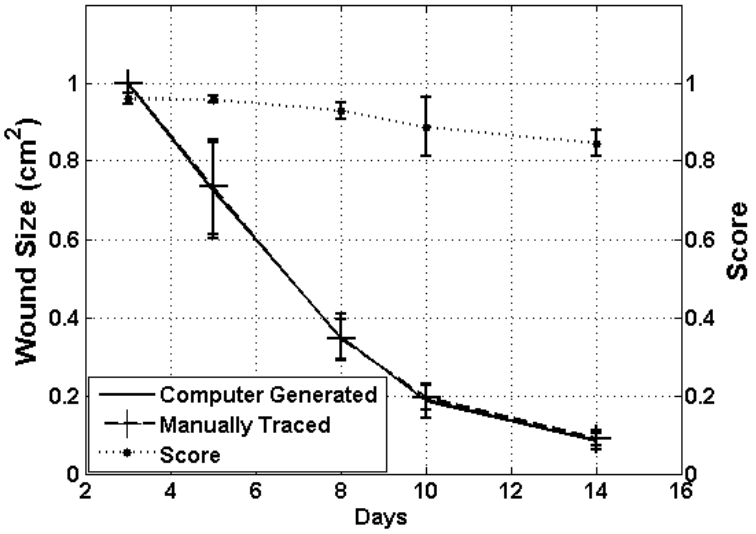

Figure 3.8 Wound healing rate of STZ rats with wound size data shown as normalized wound size. The wound healing curve from the calculated wound size using the combined image (solid black line) is very close to the manually traced wound size curve (dashed black line with "+" marker). Wound size data in square centimeters are shown on the left y-axis, and time is number of days after wound infliction. The corresponding scores for the calculated combined images (dotted line) are shown on the same plot (right y-axis).

3.3.2 Image Analysis of Wounds from Animal Model

We applied our algorithm first to wound images from an animal wound healing study. In our case, the animal images represent the ideal situation of wound images: both the skin and the wound are uniform in color, there is good color contrast between skin and wound, and images are well illuminated. Wound healing data from 16 diabetic STZ rats, from day 3 to day 14 after wound surgery, are shown in Figure 3.8 where all wound size data are normalized to a range from 0 to 1 by dividing them by the day 3 wound size. Corresponding average evaluation scores on each day are plotted above the healing curve. All evaluation scores are above 0.8 and the average score of all 56 animal images is 0.93±0.05, as shown in Figure 3.7. However, the scores tend to decrease as the wound size decreases. One reason for this is that these small wounds are nearly healed and are almost totally covered by new skin, which reduces the color contrast between the wound and the surrounding skin. Another explanation is that small objects (small wounds) are more sensitive to errors on edges than large objects because a small round area has a greater circumference/area ratio than a larger round area. The error bars on the normalized healing curves come mainly from the variation of healing rate of individual rats and are not necessarily related to the quality of image analysis. This fact is evident at day 5 when the error bars on both healing rate curves (computer-generated and manually traced) are large, while the evaluation scores at day 5 are well above 0.90.

3.3.3 Image Analysis of Wounds from Human Patients

Images of human chronic wounds are usually difficult to be recognized by a computer program because of the variability of the wound bed and the inhomogeneity of the patient skin. As discussed above, a single Green or Blue channel was adequate to analyze wound images from an animal study; however, it failed to give acceptable results for images of human chronic wounds. The wound bed often consists of red granulation tissue, yellow slough tissue, and black necrotic tissue (Mekkes and Westerhof 1995; Lait and Smith 1998; Schultz, Sibbald et al. 2003).

Skin color is different from patient to patient, and does vary from spot to spot even in the same patient. The new skin forming from re-epithelialization at the edges of the wound appears lighter in color compared to the existing old skin because less melanin is present in the new skin compared to the old skin (Sowemimo, Naim et al. 1982; Hirobe 1983; Mehendale and Martin 2001). Scars and callus around the wound will also add more color variations to the wound images. Moreover, the tissue color along the entire circumference of the wound might be quite different as we can see in Figure 3.9. Therefore, we combined information from multiple image channels to improve the scores and optimize the data analysis.

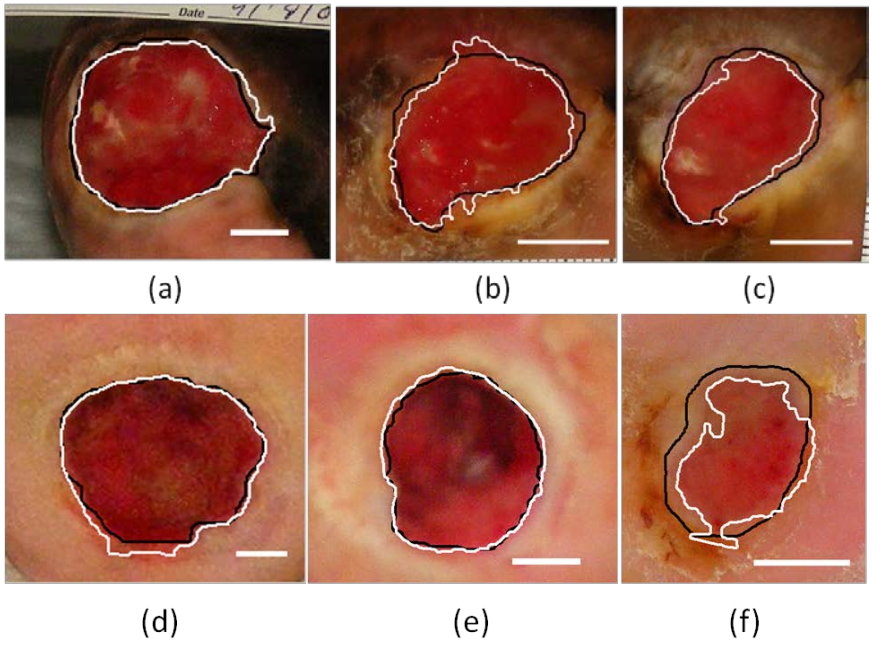

Figure 3.9 (a-c): Combined images of chronic diabetic wounds from an African-American patient on days (a) 7, (b) 42, and (c) 53 of the study. Black solid lines are manually traced wound boundaries; white solid lines are calculated wound boundaries with c = 0.5. Manually traced wound sizes on days 7, 42, and 53 were 7.61 cm^2, 2.69 cm^2, and 2.75 cm^2, respectively, while calculated wound sizes were 7.52 cm^2, 2.62 cm^2, and 2.39 cm^2, respectively, and the scores were 0.93, 0.87, and 0.85. (d-f): Combined images of chronic diabetic wounds from a Caucasian patient on days (d) 172, (e) 273, and (f) 349 of the study. Manually traced wound sizes for these days were 9.86 cm^2, 4.69 cm^2, and 2.10 cm^2, while the calculated wound sizes were 10.42 cm^2, 4.99 cm^2, and 1.57 cm^2 respectively. The scores for these three images were 0.92, 0.93, and 0.65. The white scale bar in each image represents 1cm.

The histogram of evaluation scores for all human images is shown in Figure 3.7. The maximum value of the score is 1 corresponding to 100% agreement in area and shape. While the majority of images have scores between 0.80 and 0.95, 26 images out of 99 images from human patients have a score lower than 0.8, mainly corresponding to small wounds, the presence of scars or callus with yellowish or dark color around the wounds, shadows in deep wounds, or newly developed skin. It is important to note that a score of 0.8 represents a fit about 90%, since the score is defined as the product of two ratios of the fitted area (see equation (4)).

Examples of calculated healing curves from 3 human subjects are shown in Figure 3.10 (a-c). Manually traced wound size data are plotted together with computer-generated wound size data. Computer-generated wound sizes are very close to controls (manually traced), with most scores greater than 0.8. The main contributing factor for images resulting in low scores is the complexity of the tissue in a chronic wound, which increases the inaccuracy of both manually-traced boundaries and computer-generated boundaries. Similar to our results with the animal images, a general observation in Figure 3.10 is that the low score data are most often correlated with small wound size. However, even with the presence of some low score data, the overall trend line follows very closely the trend line of the manually traced wounds.

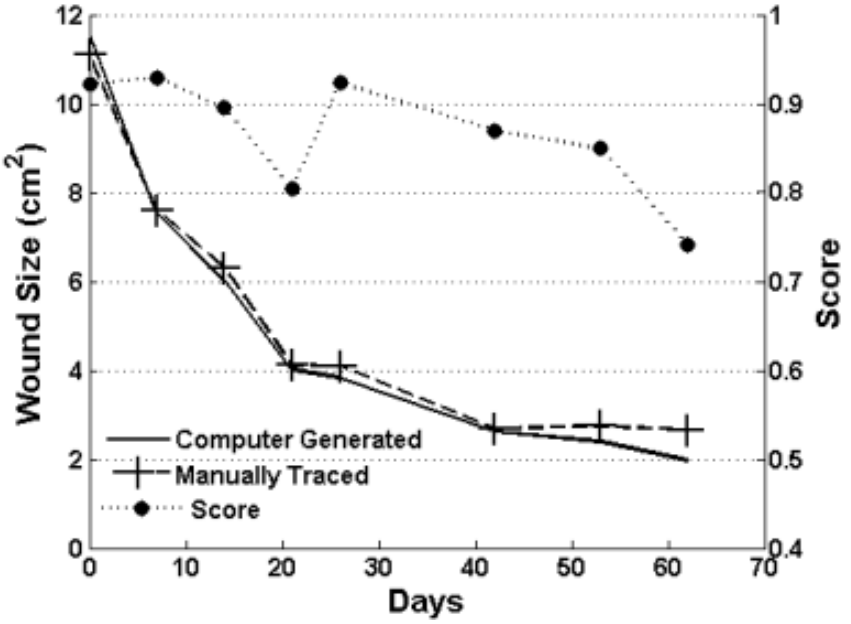

Figure 3.10 (a)

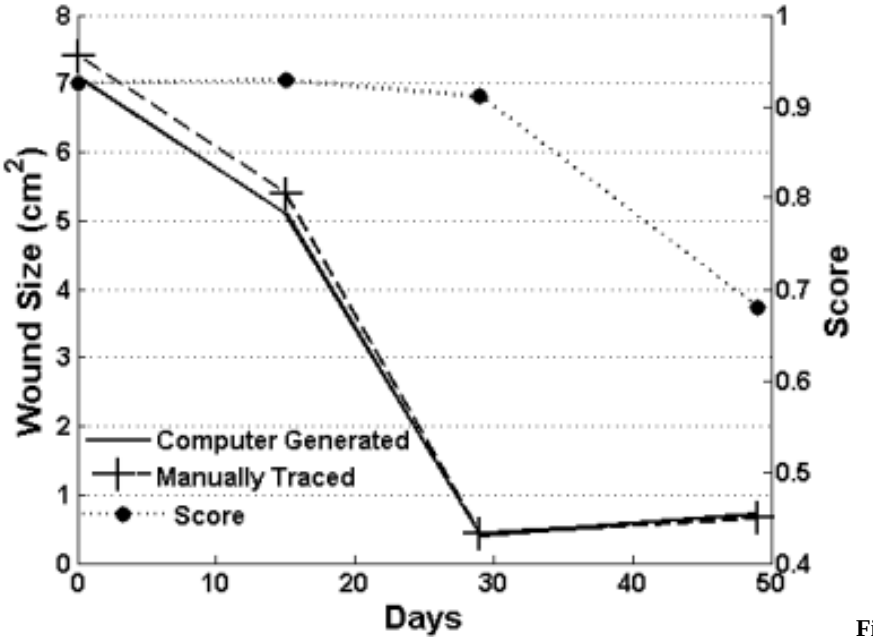

Figure 3.10 (b)

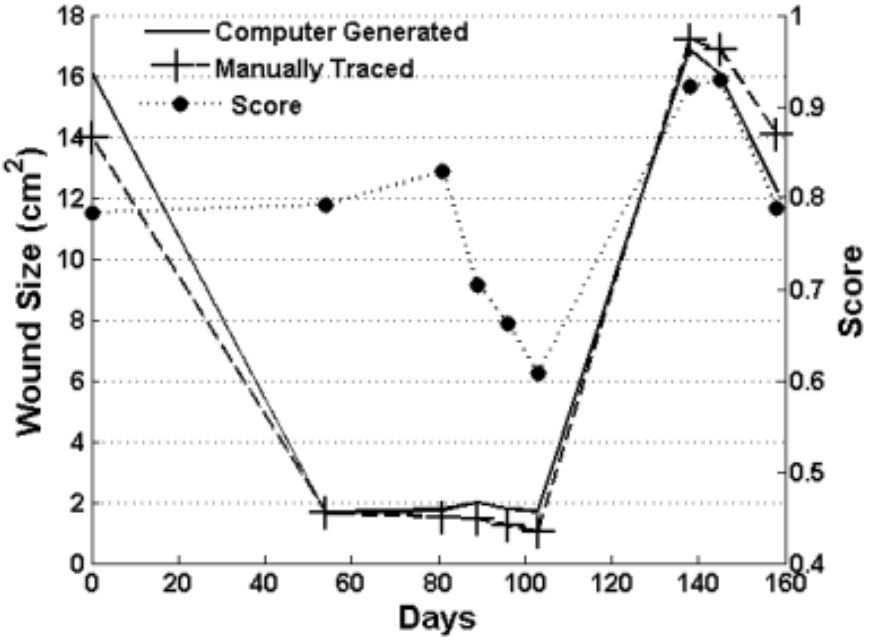

Figure 3.10 (c)

Figure 3.10 (a-c): Change of wound size with time, data from 3 diabetic patients. Calculated wound size data (solid black lines) are close to manually traced wound size data (dashed black lines with "+" marker). X axis represents time in days after the date of taking the first image. Wound size data are shown in square centimeters on the left y-axis. The corresponding scores for the calculated combined images (dotted line with "·" marker) are shown on the same plot (right y axis).

3.4 Discussion

Identification of an algorithm for detection and recognition of wound boundaries is a difficult aim to achieve because conditions and shapes of wounds are often complex. They are not only beyond the ability of a simple program to process and analyze, they are also difficult to identify even by trained human eyes. Studies have shown that despite the high inter-observer correlation of manually traced wound size among experienced medical doctors and nurses, variability did exist (Thawer, Houghton et al. 2002; Quan, Lazarus et al. 2007). However, there are two issues common to these studies. One is the method of evaluating the repeatability of wound area found by different observers (raters). A simple comparison of wound size is not adequate because it gives no information about the percentage of the area that was common to

both observers. Another issue is that none of these studies provided a correlation between the appearance of the wound bed and the repeatability in wound area among different observers. For simple cases, when wound beds have large color contrast with surrounding tissues, high repeatability is expected; when wound beds are not clean, inflammation is prevalent, and color interferences are present, the repeatability would be reduced. On one hand, this imprecision adds difficulty to the comparison of computer-generated results with results from human traced wound areas; on the other hand, it demonstrates a real need for consistent and precise wound area recognition by computer programs.

In this study, a methodology of image analysis to identify a wound boundary and calculate the wound surface area was developed and validated in wounds of both animal models and patients. The method does not require an expensive camera or proprietary codes. It is based on a point-and-shoot digital camera with images in the commonly used JPEG file format and a low cost polarizing filter and can be used in lighting conditions typical of a clinical room or animal lab and a MATLAB based program, without any fitting coefficients. Variations in skin color have been included in our study and the method works well in fair or darker skin tones. We have demonstrated that our wound boundary detection method can recognize wound area with high precision and can give us a consistent trend of wound size change with time. However, since our detection algorithm is based only on color analysis, it is not yet suitable for wounds that are nearly healed (in our study, wounds less than 1 cm^2), in which the color contrast between wound and skin is low, or deep wounds that have shadows on their images. For nearly healed wounds tracing the boundary is tricky both in the transparency tracing and our method, hence variations are larger. Another potential issue is that our current algorithm treats the wound as an average environment when calculating the boundary location. However, it is quite common that the skin and the wound bed are slightly different in color along the perimeter of the wound boundary. If intra-wound color differences need to be taken into account an improved algorithm could be

developed which accounts for variations in wound and skin color within the same image by segmenting it into wound sub-regions. Further improvements could include texture differences between wound and surrounding skin to complement our color difference and contrast based algorithm.

List of Abbreviations (abbreviations appeared in text, not in equations)

CI	Combined image
NWR	Non-wound Reference
PCC	Pixel Color Compare
ROI	Region of Interest
S	Score
SVM	Supper Vector Machine
WR	Wound Reference
TL	Threshold Level
STZ	Streptozotocin (The compound used to produce Type I diabetic rats in this study)

Chapter 4: Apply Raman Spectroscopy to Study Diabetic Foot Ulcers

In this chapter, we utilize near-infrared Raman spectroscopy to study tissue samples from diabetic foot ulcers on a small cohort of patients. In Section 4.1, the background for Raman spectroscopy and the current application of Raman spectroscopy in biomedical research is introduced. In section 4.2, patient recruiting, sample preparation, the Raman instrument and spectral pre-processing is described. Individual and averaged Raman spectra are showed in Section 4.3, and band assignments for all major bands are also given in this section. Analysis on the group averaged spectra shows that spectra collected from healing samples have relatively high intensity in bands associated with collagen and other proteins, while spectra collected from non-healing samples have relatively high intensity in bands associated with red blood cells. In section 4.4, discussions on band assignment and other experimental factors are given.

4.1 Introduction

Raman Spectroscopy is a vibrational spectroscopic method which can provide detailed information on chemical composition of the sample under investigation. Raman scattering was first discovered in the early 1900's by Krishna and Raman. However, in the first a few decades its application was impeded because of two fundamental problems: high fluorescence signal and low Raman scattering signal. Since the 1980's, with the advances of technologies such as near infrared lasers, Fourier transform technique, and charge-coupled devices (CCDs), the use of Raman spectroscopy as a characterization tool became more popular (McCreery 2000).

Raman scattering (also called Raman Effect) is illustrated in Figure 4.1 together with several other spectroscopic transitions. As a molecule is being struck by a photon with energy of $h\upsilon_0$ and transits to a short lived "virtual state" (not necessarily a true quantum state), this molecule can either return to its ground state releasing a photon with the same frequency (υ_0) or return to a higher vibrational state releasing a photon with slightly shifted frequency ($\Delta\upsilon = \upsilon_0 - \upsilon_1$). The first type of scattering event is called Rayleigh scattering, which happens primarily when the scatter is much smaller than the wavelength of the incident photon. The second type of scattering event is called Raman scattering, or more specifically, Stokes Raman scattering. Raman scattering happens only with a 10^{-7} chance compare to Rayleigh scattering (Ball 2001). If the molecule is being excited to a virtual state from a vibrational state and then returns back to the ground state, then the frequency of a released photon is also shifted from the original frequency but with a positive shift in frequency. This type of scattering is called anti-Stokes Raman scattering, but it happens with a much lower probability when compared to Stokes Raman scattering and is not involved in this study.

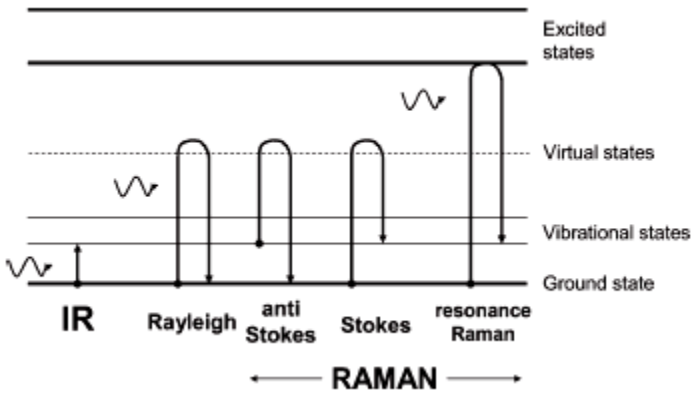

Figure 4.1 Spectroscopic transitions in several types of vibrational spectroscopy. (Figure is cited from (Ellis and Goodacre 2006)).

A Raman spectrum is a plot of signal intensities over a range of frequencies expressed in wavenumbers (cm^{-1}).Each peak corresponds to a Raman shift in frequency from the original incident light, while the shift of light energy ($h\upsilon_0$- $h\upsilon_1$) equals to the energy difference between a vibrational state and the ground state. This vibrational state corresponds to a specific molecular vibrational mode. Theoretically, a molecule with N atoms can have 3N-6 or 3N-5 vibrational degrees of freedom, corresponding to non-linear or linear molecules. And we can imagine even for a small biological molecule, which usually has a molecular weight of a few hundred, its vibrational spectrum can be rather complicated. Vibration modes can be generalized as stretching modes (υ) and bending modes (δ). The stretching vibration involves the change of bond length, while the bending vibration (also called deformation) involves the change of bond angels. Since a molecule can usually have multiple Raman active vibrational modes, a Raman spectrum of the molecule will be featured with multiple peaks/bands at specific Raman shifts (cm^{-1}) corresponding to these vibrational modes. In Figure 4.2, a Raman spectrum of cholesterol is shown as an example.

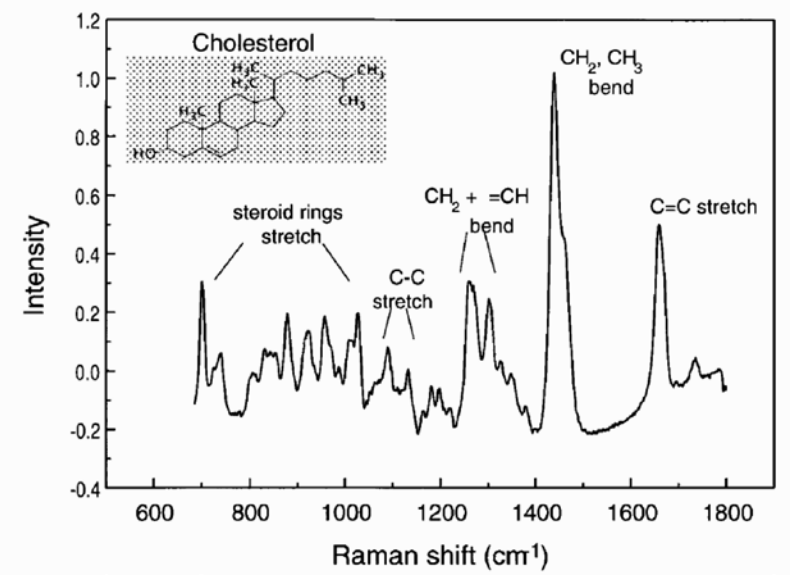

Figure 4.2 Raman spectrum of cholesterol with typical vibrational bands labeled. Background is removed by fourth order polynomial fit to the raw spectrum. The Figure is cited from (Hanlon, Manoharan et al. 2000).

Raman spectroscopy has increasingly gained more attention in biomedical studies due to its potential as a non-destructive and non-invasive diagnostic tool (Hanlon, Manoharan et al. 2000). As mentioned before, a Raman spectrum is an intensity plot over a series of vibration frequencies, where the frequency corresponds to a particular molecular vibration mode and signal intensity corresponds to the abundance of a molecular structure. A prominent benefit of using Raman method it is that the presence of water in sample does not produce strong water signal in Raman as it would in Near Infrared (IR) spectroscopy. This feature gives Raman spectroscopy the possibility of measuring biological samples without the drying process that could denature the sample or measuring samples non-destructively. Meanwhile, it has been shown that Raman devices can be coupled with optical fibers for in vivo measurements, which makes the device more appealing in terms of clinical purpose (Shim and Wilson 1997). So far, Raman spectroscopy has been used to examine tumor tissue in lungs (Huang, McWilliams et al. 2003), esophagus (Shetty, Kendall et al. 2006), stomachs (Teh, Zheng et al. 2010), colons (Beljebbar, Bouche et al. 2009), bladders (Crow, Molckovsky et al. 2005); blood samples (Wood, Tait et al. 2001; Deng,

Wei et al. 2005), serum (Pichardo-Molina, Frausto-Reyes et al. 2007), body fluid (Sikirzhytski, Virkler et al. 2010) and single cells (Krafft, Knetschke et al. 2003; Huang, Griffiths et al. 2004). Crane et al. published their pilot study on monitoring combat wounds with Raman spectroscopy (Crane, Brown et al. 2010). Han et al. analyzed serum from diabetic patients with surface enhanced Raman spectroscopy (SERS) (Han, Yan et al. 2009). However, to the best of our knowledge the possibility of using Raman as a tool to study diabetic foot ulcers has not been previously investigated.

Despite the advantages of using Raman spectroscopy in biomedical studies, Raman spectra from biological samples are usually not easy to analyze due to the complexity in their constituents. As a result, these Raman spectra are often seen as unresolved continuous spectral profiles; weak bands could become indiscernible bumps on a neighboring strong band; neighboring Raman bands may stack up and result in a new peak position; multiple vibration modes in different molecules may produce bands at the same frequency; and subtle changes of the surrounding environment of a molecule can lead to a shift of Raman bands. All these factors present difficulties in the analysis of the spectra in a traditional way, making it nearly impossible to identify every single featured band and assign it to a particular molecular structure or component in the sample.

A widely adopted approach of analyzing Raman spectral data is using multivariate statistical methods to extract information from a large amount of data or to cluster spectral data. Principle Component Analysis (PCA) is the most commonly applied method to reduce the data size (Chatfield and Collins 1980; Hanlon, Manoharan et al. 2000). Factor analysis (FA) has also been utilized to extract tissue component information (Omberg, Osborn et al. 2002; Crane, Brown et al. 2010). The use of classification or clustering methods such as linear discriminate analysis (LDA) (Koljenovic, Choo-Smith et al. 2002; Teh, Zhene et al. 2010), cluster analysis (Zhang, Henson et al. 2005), hierarchical cluster analysis (Beljebbar, Bouche et al. 2009), and neural

networks (Gniadecka, Philipsen et al. 2004) have been reported. Supervised classification using SVM has also been attempted to classify human nails (Widjaja, Lim et al. 2008).

In the current chapter, we investigate the possibility of using Raman spectroscopy as a tool to differentiate between healing and non-healing diabetic foot ulcers. Samples of wound tissue were collected during surgical debridement from open wounds on diabetic patients. Raman spectra were collected analyzed. Our analysis shows spectra collected from healing wounds show relatively higher intensities at bands from collagen.

4.2 Materials and Methods

4.2.1 Patients and Sample Collection

All patients were recruited into this study from Drexel University Wound Healing Center (Philadelphia, PA). These patients had been diagnosed with diabetes mellitus and had an open wound on their foot, with the wound area being greater than 1 cm^2 at the beginning of the study. A picture of a diabetic foot ulcer is showed in Figure 4.3 (a) as an example. On their weekly or bi-weekly visits to the clinic, the patients received wound care including wound cleansing, debridement, offloading and dressings. A wound was categorized as "healing" if it closed within 20 weeks of visiting the clinic; otherwise, it was categorized as a "non-healing" wound. A more detailed description of patient recruitment can be found in a previous paper on human subject diabetic foot ulcer study (Weingarten, Neidrauer et al. 2010). In total, 23 samples were collected from 7 wounds (6 patients). Among these samples, 14 were collected from healing wounds and 9 samples were collected from non-healing wounds. The study protocol was reviewed and approved by the Drexel University College of Medicine Institutional Review Board.

Figure 4.3 (a) An example picture of a diabetic foot ulcer (left). (b) An example image of cryosectioned tissue slide under a microscope (right). The scale bar in the image is 500 μm.

4.2.2 Sample Preparation

Wound debridement samples were harvested during debridement by the wound clinician, immediately placed on dry ice, and then transferred to small vials and stored at -80˚C without further processing until use. Samples were embedded in the optimal cutting temperature compound (Tissue-Tek® 4583 CRYO-OCT Compound, Sakura Finetek, Torrance, CA). Frozen sample cubes were then sectioned using a cryotome into 25 μm slices at -20˚Cand placed onto quartz slides. Slides with tissue sections were dried in air before continuing on with Raman Spectroscopy. From each sample, two to four 25 μm sections were collected at an interval of 0.1 mm, depending on the size of the tissue sample. Collecting multiple sectioned slices was to ensure that Raman spectra were obtained from different locations on each tissue slice, accounting for variability within each tissue sample. Tissue slides were then brought to Raman characterization without further treatment.

4.2.3 Raman Instrumentation

A LabRAM HR800 confocal Raman microscope system (Horiba Jobin Yvon, Edison, NJ, USA) was used in this study. The Raman microscope system is integrated with an Olympus BX41 microscope, and is equipped with a *xyz*-motorized stage, with mechanical resolution of 0.1

μm in all directions. The 785 nm laser was illuminated through a 100×/0.9 objective lens, focusing on the sample to get a maximum Raman signal. Laser power on the sample was no greater than 15mW. The laser beam has a diameter of 15 μm. Confocal pinhole was set to 300 μm for all measurements. Raman signal was collected by a Peltier air-cooled CCD detector that has 1024×256 pixel resolution. Spectra were collected from 600 – 1800 cm^{-1} using a 600 line/mm grating unit, giving the spectra a resolution of 2.5 cm^{-1}. The system was calibrated prior to the beginning of this study, and the Raman band of silicon wafer at 520 cm^{-1} was confirmed before every measurement on a tissue slide. Raman signal collecting time was 10 or 20 seconds, and the number of accumulations was set to 2 to accommodate the algorithm for removing cosmic spikes. Cosmic spikes were removed by the system software (LabRAM, Horiba Jobin Yvon, Edison, NJ, USA) using the "multi accum" mode, by which spikes that were not detected at the same wavenumber were considered cosmic ray signal and were removed from spectra. Raman spectra were collected from each sample slide on multiple sites, which were about 100 μm apart from each other. The overall number of spectra from each sample ranges from one hundred to a few hundred depending on the actual size of tissue sample.

4.2.4 Spectra Pre-processing

All spectra processing were done in MATLAB R2010b (The Mathworks Inc., Natick, MA) using in-house written code. Spectra were smoothed using 5-points moving average filter. Baseline correction was done using a method described by Lieber and Mahadevan-Jansen (LMJ method) (Lieber and Mahadevan-Jansen 2003). In the LMJ method, each spectrum is fitted with a 5^{th} order polynomial to find a first baseline; a spectra remnant, consisting of the baseline and the Raman spectrum that is below the baseline, was used for another round of 5^{th} order polynomial fitting, yielding a new baseline. This process is repeated until the least square residual between the current and the last baseline is less than a predefined threshold value (0.4 was used here, giving a visually acceptable fitted baseline). Subtracting the baseline from the original spectrum

gives the LMJ-corrected spectrum. All spectra were normalized by adjusting the area under spectra to the same value. Spectra were additionally aligned by centering the phenylalanine ring-breathing band to 1004 cm^{-1}; this band may show slight shifts due to instrument variation during the time of collecting spectra. Spectra were then trimmed so that only common frequencies were kept. After pre-processing, the data is represented as a spectra matrix D_0, with each column corresponding to signal intensities at a frequency and each row representing a Raman spectrum for a sample. Data set D_0 consisted of a total of 3037 rows (spectra) and 1767 columns (Raman shift).

4.3 Results

A summary of the number of samples and the number of spectra collected for this study is given in Table 4.1. As there were multiple samples collected from one wound at multiple clinical visits, we ordered these samples by time.

Table 4.1 Summarize of the number of samples and the number of spectra collected.

Group	Subject ID	Number of Spectra Collected from Each Sample							
		1st (visit)	2nd	3rd	4th	5th	Summary		
Healing (n=14)	**s39b**			51	96	43	190	**1824**	**3037**
	s6		193	130	273	151	747		
	s34					77	77		
	s41				231	135	366		
	s45		117	141	107	79	444		
Non-healing (n=9)	**s39**		82	21	50	129	282	**1213**	
	s32	158	240	57	264	212	931		

All Raman spectra from both healing and non-healing samples are shown in Figure 4.4(a), and averaged spectral profiles of each group are also shown in Figure 4.4(b). The difference spectrum showed in Figure 4.5 is produced by subtracting the average spectrum of non-healing samples from the average spectrum of healing samples. Spectra from both groups show almost the same spectrum profile with only slight difference in intensities on several frequencies. Some of the major bands can be easily assigned to vibration modes in lipids, proteins, and carbohydrates. The prominent band at 1004 cm^{-1} can be assigned to Phenyl ring breathing (ν(C-C)) from phenylalanine, band at 1250 cm^{-1} can be assigned to amide III, band at 1341cm^{-1}can be assigned CH_3 or CH_2twisting model of proteins and nucleic acids, the band at 1449 cm^{-1} can be assigned to δ(CH_2) and δ(CH_3) scissoring in proteins, and band at 1658 cm^{-1} can be assigned to ν(C=O) of amide I from α-helix of proteins. Other band assignments are summarized in Table 4.2. For a more detailed description of Raman band assignments, please refer to Movasaghi's review (Movasaghi, Rehman et al. 2007).

Table 4.2 Major peaks on averaged Raman spectra from healing and non-healing wounds.

Frequency (cm^{-1})	Vibration Mode	Tentative Molecular Assignment	Reference
622	δ (C-C) twisting	Phenylalaine (Phe) of proteins	a
645	δ (C-C) twisting	Phenylalaine or tyrosine,	d
665	Ring deformation	Pyrrole, Observed in RBC	b
	ν (C-S)	Cystine	d
	Ring breathing	G, T ring DNA bases, tyrosine-G backbone in RNA	e
677	Ring breathing	DNA bases	e
702	Sterol ring stretching	Cholesterol or cholestor esters	a
720		Choline, or DNA	g
756	ν (pyrrole ring breathing)	Red blood cells	b
		Tyrosine, Tryptophan	g
829		Proline, hydroxyproline, tyrosine, ν(PO_2^-) stretch of nucleic acids	d
854	ν (C-C); (C-CH) ring breathing	Proline, tyrosine, Polysaccharide	c
	Ring breathing or stretching	Proline, hydroxyproline, tyrosine	d
880	ν (C-C)	Hydroxyproline in collagen	i
	δ (ring)	Tryptophan	g
899	(C-O-C) skeletal mode	Monosaccharides (β-glucose)	g
		Deoxyribose	g
	C-C skeletal backbone	Protein	e
936	ν (C-C)	Proline, hydroxyprolien, skeletal of collagen backbone	d
961	ν (PO_4^{3-})	Hydroxyapatite, carotenoid, cholesterol	g
970		Phosphate monoester groups of phosphorylated, proteins & cellular nucleic acids	g
	δ (Pyr deform)$_{asym}$ and/or γ (=C_bH_2)$_{sym}$	Observed in RBC	b
	ν (C-C) wagging, or C-C backbone of collagen	Collagen	g
1004	ν (C-C) Phenyl ring breathing;	Phenylalanine in proteins	g
1033	ν (C-H), in-plane	Phenylalanine , (Protein/Collagen)	d
	δ (C-H)	Phenylalanine	c
1128	ν (C-C) skeletal of acyl backbone in lipid (trans conformation)	Lipids	d
	ν (C-N)	Protein	e
1157	ν (C-C) or ν (C-N)	Protein	g
1174	δ (C-H) bending	Tyrosine (collagen type I, protein), phenylalaine, or guanine.	g
	ν (pyrrole half-ring)$_{asym}$	Observed in RBC	b

1212	ν (C-C_6H_5)	Tryptophan and phenylalanine	h
	δ (C-H)	Observed in RBC	b
1250		Amide III band	d
	ν (C-N) and N-H bending	Amide III band of proteins	g
1306	δ (C-H)	Observed in RBC	b
	CH_3/CH_2 twisting or bending mode	Lipid/collagen	d
1318		Guanine (B-Z marker or ring breathing), or Amide III in α-helix.	g
		Adenine	f
1341	CH_3/CH_2 twisting	Proteins and nucleic acids	j
1392	CH_2 deformation	(Not from collagen)	d
	δ (pyrrolequater-ring)	Observed in RBC	b
1449	δ (CH_2) and δ (CH_3)	Proteins and lipids	c
	CH_2/CH_3 deformation		d
1561	ν (C-C) of heme	Observed in RBC	b
1584	ν (C-C) of heme	Observed in RBC	b
1605	ν (C-C) of heme	Observed in RBC	b
	ν (C=C)	Phenylalaine or tyrosine	e
1620	ν (C=C) of heme	Observed in RBC	b
1658	ν (C=O)	Amide I, alpha-helix of proteins	c
1732	ν (C=O)	Cholestrol esters	a

Notations: (i) ν, stretching; δ, in-plane deformation; γ, out-of-plane deformation; RBC, red blood cell. (ii) Band assignment are find in these references a: (Manoharan, Baraga et al. 1992), b (Wood, Tait et al. 2001), c: (Huang, McWilliams et al. 2003), d: (Cheng, Liu et al. 2005), e: (Chan, Taylor et al. 2006), f: (Pichardo-Molina, Frausto-Reyes et al. 2007), g: (Movasaghi, Rehman et al. 2007), h: (Teh, Zhene et al. 2010), i: (Ikoma, Kobayashi et al. 2003), j: (Teh, Zheng et al. 2008).

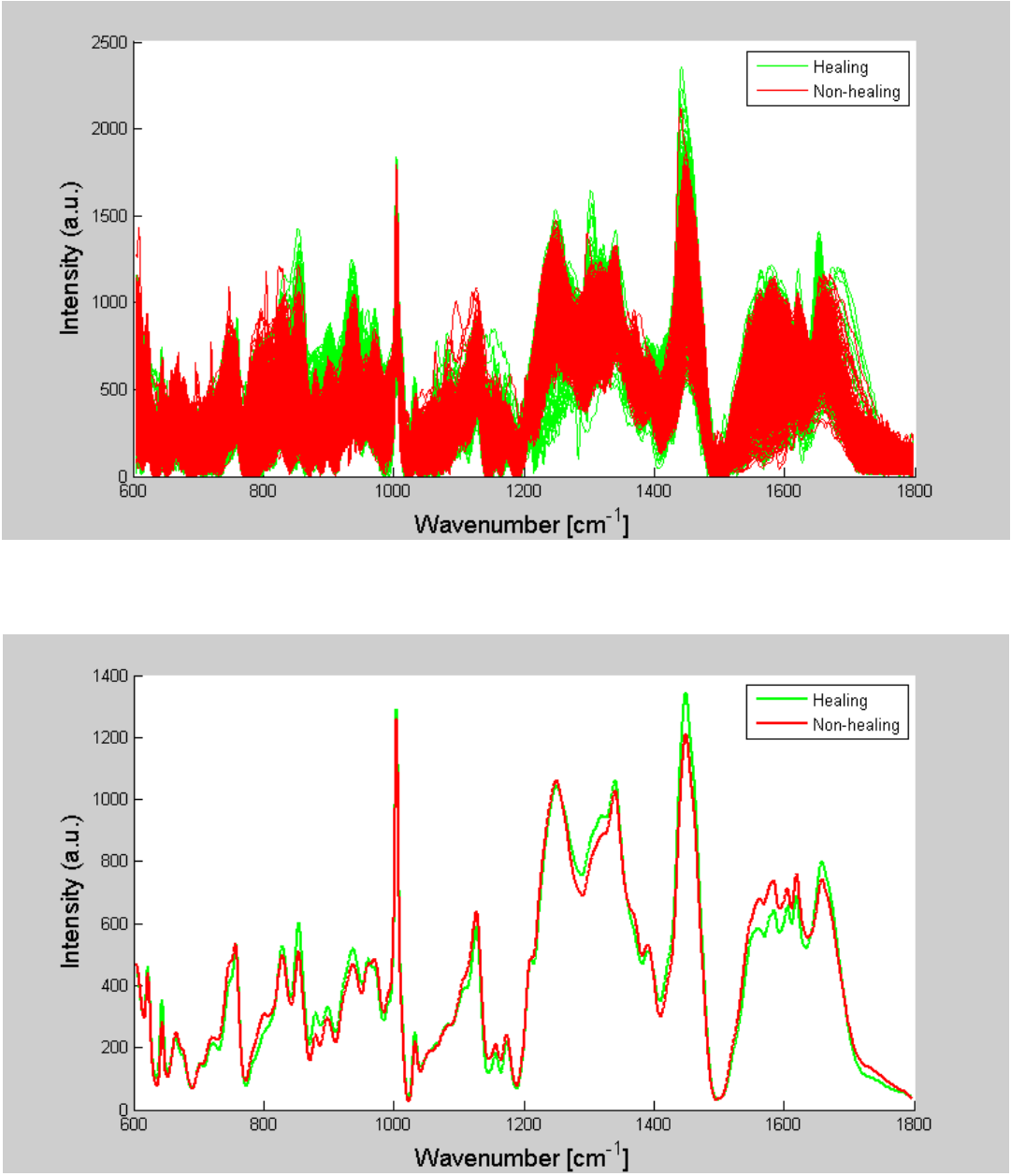

Figure 4.4 Individual Raman spectra from both healing and non-healing samples (a, top) and the averaged spectra of each group (b, bottom). Spectra from healing samples are plotted in green solid lines, while spectra from non-healing samples are plotted in red solid lines.

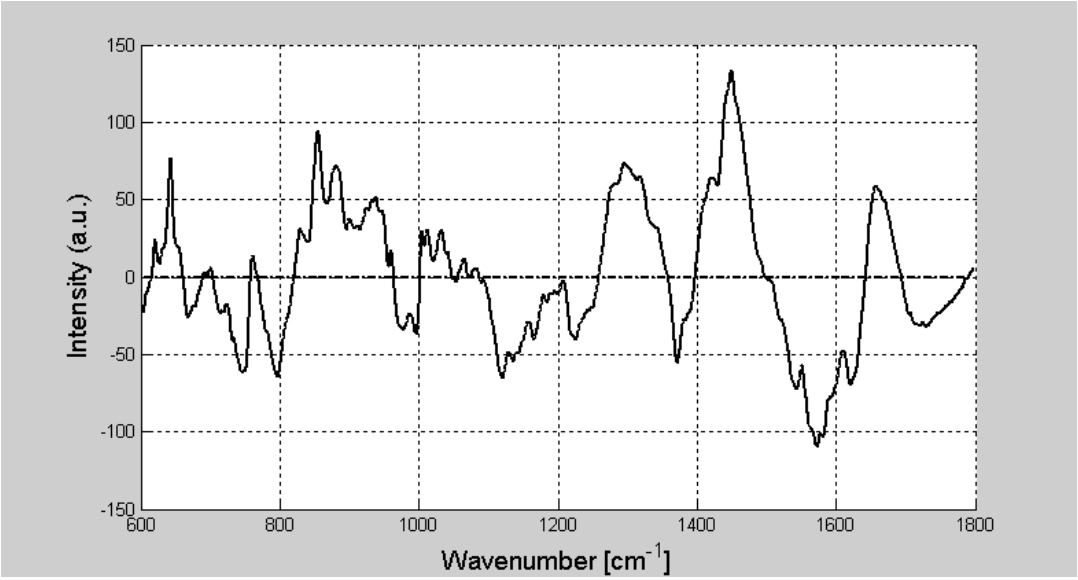

Figure 4.5 A plot of spectral difference calculated by subtracting the average non-healing spectra from the average healing spectra. The positive difference (where the curve is above the central line) indicates the mean spectrum of healing samples are higher in intensity compare to the mean spectrum of non-healing samples, and vice versa.

4.4 Discussion

4.4.1 About Raman Spectra Analysis

Raman spectra from both groups are very similar. Comparing the averaged spectra from healing and non-healing samples to each other (Figure 4.4(b)), all major bands are common to both groups and intensities of all frequencies are very close. This is due to two reasons. Firstly, the reality that samples we collected from both groups are essentially the same type of tissue and they share high similarity. Secondly, patients in both groups have diabetes and so healing processes are both impaired to some extent. While there is proof of certain substances are higher or lower in content in impaired-healing wounds than normal-healing wounds, it is difficult to imagine that difference is great enough to show as a distinct feature on Raman spectra.

Table 4.3 The band assignment for major spectral differences between the healing and non-healing group.

Difference	Wavenumber (cm^{-1})	Assignment	
		Vibration Mode	Molecule
Intensity is high in healing group	645	(C-C) twisting	Proteins
	854	(C-C) stretching	Collagen/other
	876	(C-C) stretching	Collagen
	936		
	1320	Guanine, or Amide III	Nucleic acids, protein/lipids
	1451	CH_2/CH_3 twisting	Proteins/Lipids
	1658	Amide I	Proteins
Intensity is high in non-healing group	665	Pyrrole ring deformation	Red blood Cells
	756	Pyrrole ring breathing	
	796		
	971	Pyrrole ring deformation	
	996	(C-C) stretching	
	1174	Pyrrole half ring stretching	
	1212	(C-H) twisting	
	1561	Porphyrin in-plane vibration	
	1582		
	1605		
	1618		

Despite the high similarity between the averaged spectra from two groups, we noticed differences (Figure 4.5). And some of the major band intensity differences are summarized in Table 4.3 (Please refer back to Table 4.2 for references). The healing group has higher intensity than non-healing group at frequency of 645, 854, 876, 936, 1320, 1451, and 1658 cm^{-1}; also note that intensities in healing group are lower than non-healing group at frequencies of 1561, 1582, 1605, 1618 and 1732 cm^{-1}. The band at 645 cm^{-1} can be assigned to phenylanine or tyrosine in proteins (Cheng, Liu et al. 2005; Movasaghi, Rehman et al. 2007). Both phenylalaine and tyrosine are amino acids presented in either collagen or other proteins in muscle and plasma. It was reported by Reeds et al that there was an increase in concentrations of phenylalaine in acute-phase proteins (Reeds, Fjeld et al. 1994). Askanazi et al. also confirmed elevated phenylalaine

and tyrosine content in muscle and plasma associated with operative trauma (Askanazi, Furst et al. 1980). The high band signal at 645 cm^{-1} could come from either the collagen in newly formed ECM or from proteins in muscle tissue. The peak at 876 cm^{-1} was assigned to proline or hydroxyproline, and had been used as an indicator of collagen (Teh, Zheng et al. 2010). Difference at the intensity of 876 cm^{-1} suggests low collagen content in non-healing wounds on average. This is consistent with our knowledge about impaired diabetic foot ulcer where people have found high level of MMP-9, a potent proteases that can degrade collagen as well as a sign of inflammation (Trengove, Stacey et al. 1999; Falanga 2005; Liu, Min et al. 2009; Widgerow 2011). Similarly, bands at 854 and 936 cm-1 are also indicators of low collagen content in non-healing wounds (Cheng, Liu et al. 2005; Movasaghi, Rehman et al. 2007). The band at 854 cm^{-1} is reported to be assigned to ring breathing mode of proline, hydroxyproline or tyrosine. Given the general occurrence of tyrosine in tissues is low (only 3% in collagens, 5% as free amino acids in plasma), the source of the band is more likely from proline and hydroxyproline in collagen. The peak at 1658 cm^{-1} is assigned to amide I band, and peak 1451 cm^{-1} is assigned to CH_2/CH_3 deformation of collagen or lipids (Movasaghi, Rehman et al. 2007). These bands (854, 876, 936, 1451, and 1658 cm^{-1}) collaboratively show a high collagen content in healing wounds compare to non-healing wounds. While the shoulder band at 1320 can be assigned to ring breathing mode of nucleic acids or CH_2/CH_3 twisting mode of collagen/lipids (Movasaghi, Rehman et al. 2007). The higher nucleic acids presence in wound bed suggests a possible higher count of cells in this region.

On the other hand, intensities of some bands are high in non-healing group as it can be observed in Figure 4.4(b) as well as in the difference spectrum in Figure 4.5. The 1732 cm^{-1} band is assigned to ester bond in lipids (Movasaghi, Rehman et al. 2007). Peaks such as 1561, 1582, 1605, and 1618 cm^{-1} can be assigned to red blood cells (Wood, Tait et al. 2001). Other bands that associated with red blood cells, such as 668, 756, 796, 971,996, 1174, and 1212 cm^{-1} are all slightly higher in intensities in non-healing spectrum compare to healing spectrum. Most of these

bands between 668 and 1212 cm^{-1} are associated with vibrations in pyrrole ring, which makes up porphyrin in hemoglobin. All these bands together suggest that spectra in non-healing group have more Raman signal from red blood cells.

We should note that spectra that are shown here had been normalized by adjusting the area under the entire spectrum. So the intensity of a Raman band in Figure 4.4 is in fact a relative signal intensity of one band against the entire spectrum. Therefore, the spectral difference shown in Figure 4.5 only reflects the relative band intensity difference. More specifically, the differences we can see in Figure 4.5and Table 4.3 don't necessarily mean non-healing samples have more counts of red blood cells, it only suggests red blood cells made up more percentage out of the total Raman active materials since there might be less other Raman active materials, such as collagen. However, the biological/pathological explanation of why non-healing wounds contain relatively high level of these constituents is yet to be elucidated.

Generally speaking, two mechanisms can contribute to the differences of band intensities we observed between healing and non-healing groups. One is the difference of relative content of certain molecules (concentration). The other one is the difference of the relative content of a molecule in certain molecular conformations or orientations (Snyder, Strauss et al. 1982; Teh, Zheng et al. 2008). In our case where the wound tissue sample is a mixture of many different biomolecules, it is possible that both mechanisms play their roles.

Giving definitive assignment to every Raman band is difficult. On one hand, tissues are heterogeneous in terms of biochemical composition; the single vibration mode that contributes to a Raman band can come from different biomolecules. For example, peak at 1451 cm^{-1} is assigned to CH_2 bending mode either from proteins or lipids (Movasaghi, Rehman et al. 2007). On the other hand, multiple vibration modes can contribute to the Raman intensity at a certain frequency. It is rather common to find that two Raman bands are close to each other and part of their bands is

overlapped. And so the signal in the overlapped region is the summation of signal from multiple vibration modes.

Conventionally, Raman spectra analysis was done on band intensity or band intensity ratio. Single band intensity gives information of the content of certain molecular structure relative to the entire Raman active substance, while the intensity ratio can give information about the relative content of the two molecular structures contributing to the two peaks. Some of the intensity ratios that have been investigated in other Raman spectroscopy studies will be briefly reviewed in Chapter 5. And we included both peak intensities and peak intensity ratios as feature data to train a Support Vector Machine classifier in the Chapter 5.

4.4.2 The Choice of Substrate

Many researchers chose to use CaF_2 as substrate for sectioned samples; however there are several studies mentioned using aluminized quartz or quartz as substrate for their samples. Schuster et al. compared the background signal from three substrates: glass, quartz and CaF_2. When excited with 632.8 nm laser, the glass slide has strong fluorescent background and the signal intensity gets stronger as it moves from 600 cm^{-1} to 1400 cm^{-1}; the quartz slide has better performance since it has relative lower fluorescent background and flat spectrum profile in 600-1800 cm^{-1} region; while the calcium fluoride slide has low and flat background fluorescence in fingerprint region except a sharp peak at 322 cm^{-1}, which makes it the best choice of substrates when we only interested in 600-1800 cm^{-1} region (Schuster, Reese et al. 2000). In another study, Filik et al. stated that quartz slide was not an option since it produces strong fluorescent background when excited with 830 nm laser (Filik and Stone 2008). However, Ortiz et al. showed in their work that spectra collected from proteins on either quartz or SpecRIM ® slides (made from stainless steel) had little difference under 632 nm excitation wavelength (Ortiz, Zhang et al. 2006). However, some studies did use quartz cuvettes or slides. In one study, Huang et al. used quartz substrate in their study of bacteria. A 532 nm laser was used to excite bacteria spread on

quartz slides and they and managed to discriminate cell growth phases and species (Huang, Griffiths et al. 2004). In another study, cells were directly grown on quartz slides, fixed with formalin and then Raman spectroscopic maps were collected from cells with 785 nm excitation wavelength (Krafft, Knetschke et al. 2006).

The use of CaF_2 or aluminized glass substrate would be ideal for Raman study, but quartz slides are still acceptable in this study due to several reasons. Firstly, laser beam will be focused on tissue surface through a 100x/0.9 objective lens, which greatly reduced the sampling volume of laser beam. According to the Manual of LabRAM HR800 system, the FWHM (Full Width at Half Maximum) of a 100x/0.9 objective lens is about 2 μm while the FWHM of 10x/0.25 objective lens is about 183 μm (according to the manual for LabRAM HR800 system). Secondly, in this confocal Raman system whose advantage is "considerable reduction of the depth of focus and thus an increased Z discrimination". Thus, only the light signal from the certain layer of depth will be allowed to travel back to the detecting unit, while other scattered light that was not in the confocal light cone will be blocked. In the current experiment setting (objective lens 100x/0.9, confocal pin-hole 300 μm), the FWHM is 2.5 μm according to the Manual, however the actual laser beam spot is about 15 μm in diameter when measured on by the video camera installed on the Raman system. Thirdly, given that tissue sample investigated in this study is opaque, Raman signal should come mainly from sample surface. And finally, since all tissue slides were prepared in the same way, even if background signal from quartz substrate did exist, the effect should be consistent to all collected spectra.

4.4.3 Collecting Sample from Debridement

Debridement sample is heterogeneous in its texture, while in other similar studies, biopsy samples are used. The consistent texture of the sample allows good histological stained slides being produced and allows Raman mapping to be performed. In addition, these samples are

heterogeneous in their composition. Samples were collected from surgical debridement of diabetic foot ulcers, which were comprised of some solid small pieces of tissue, scraped tissue on the surface of wound bed, tissue fluid and blood. The heterogeneous nature of the sample requires more spectra to be collected to ensure a good sampling from the entire sample. While the dimension of a typical sectioned sample is about a few millimeters, the diameter of the laser beam from our Raman system is 15 μm, which won't give us a good sampling of the whole sample in just a few randomly selected spots on the sample slide. Therefore we collected a large number of Raman spectra from each sample.

4.4.4 Sample Handling

A common methodology of studying Raman spectroscopy of biological tissues is to confirm what discovered by Raman spectra with histological stained tissue slides. Usually, one 20 or 25 μm tissue slide will be sectioned for collecting Raman spectra and an adjacent 5 μm slide will be sectioned for staining. Most tissue samples collected for Raman spectroscopic studies were preserved by snap-frozen technique, in which tissue samples are wrapped in aluminum foil or put in a small vial and then snap-froze in liquid nitrogen. This procedure ensures minimum ice crystals formation during the fast freezing process, which is critical to preserve tissue structure and produce good histology slides. However, debridement samples collected for this study were not snap-frozen in liquid nitrogen. Samples were put in small vials and kept cool on dry ice until transferred to -80 °C refrigerator. This may lead to formation of ice crystals in the sample and destroyed fine structure of tissue. However, it should also be noted that a sample collected by surgical debridement may not have a fine structure at all since it's scraped from wound surface. For future study, snap-frozen technique is recommended for sample preservation.

We should also be aware that during the sample preparing process there are factors that could affect the Raman spectrum. Firstly, macromolecules in the sample such as proteins can

undergo structural changes during freezing, drying or fixation. Faolain et al. compared Raman spectra collected from fresh tissue sample, frozen sample and fixed sample, and spectral differences due to protein structural change or chemical contamination are found (Faolain, Hunter et al. 2005). Although fixation is not adopted in this study, freezing and drying are part of sample preparation procedures in this study. Secondly, new chemicals could be introduced to the sample during sample preparation. The O.C.T. compound, which was used as embedding medium for cryosectioning, contains10.24% of polyvinyl alcohol (PVA), 4.26% polyethylene glycol (PEG) and 85.5% nonreactive ingredient (water). And both PVA and PEG have Raman-active vibration modes. Traces of PVA and PEG might be found on the broader areas of cross-sections of tissue due to OCT compound diffusing into samples or even on the cross-sections due to blades ran across the frozen sample cube can bring these molecules from other part. Attention had been paid to not collecting Raman spectra from spots close to the broader on tissue sections.

Despite of these adverse effects from freezing sample and cryosectioning, there seems no better alternative sample handling method exists. Any handling method can result in spectral differences between fresh tissue sample and processed tissue sample (Huang, McWilliams et al. 2003; Faolain, Hunter et al. 2005).

4.4.5 Sample Size

A large number of spectra have been collected; however, the sample size in this study is small: 23 debridement samples of 7 wounds from 6 patients. Among these 7 wounds, only 2 wounds are non-healing.

Chapter 5: Supervised Classifications of Raman Spectra from Diabetic Foot Ulcers

The goal of this chapter is to apply supervised classification methods to Raman spectral data. In section 5.1, a brief introduction of pattern recognition and classification is given. In section 5.2, methods that are involved in Raman spectral data classification are described. Raman spectra were categorized as healing or non-healing based on the wounds they were collected from. Both Raman intensity data and intensity ratio data are included as features for later classification. Feature selection was done by performing unpaired t-test and ranking the features by their p-value. Selected feature data is fed to a support vector machine (SVM) classifier for training, and the performance of the classifier was evaluated by 10-fold cross-validation. Section 5.3 presents the classification results. The trained SVM classifier was able to predict the category of a spectrum with 85.2% accuracy. The prediction of the category of a sample was of 95.7% accuracy when the averaged spectrum of each wound sample was fed to the trained SVM classifier. In section 5.4, an attempt of classifying wounds by the color of wounds is described. And in section 5.5, possible causes of misclassification are analyzed.

5.1 Introduction

As a result of the advances in data collection various technologies in the past few decades, we are now living in a data-rich era. In the meanwhile, modern scientific research also becomes more and more data-driven. With the increase in data size, it becomes a daunting task to conduct the data analysis using traditional methods. Novel data analysis methods are required to extract non-trivial information that can lead to new discoveries.

The goal of statistical data analysis is to recognize the underlying patterns in the data. Approaches in data analysis can be categorized either as confirmatory or exploratory, depending on whether there exists an appropriate model for the data. Both types of data analysis approaches rely on classifying the data, based on either fitting the data to a postulated model or to the naturally existed groups (clusters) revealed through analysis (Jain, Murty et al. 1999). Classification can be supervised classification (discriminant analysis) or unsupervised classification (clustering). In the supervised classification, each measurement/instance has a known label and the goal is to find out underlying patterns in the current data set in order to predict labels for new measurements. In contrast, the unsupervised classification is provided with unlabeled measurements/samples, and the goal is to group these data into meaningful clusters based on their similarity. There are four best known pattern recognition approaches, among which statistical approach is most widely used and studied. The classification method applied to analyze current Raman spectral data falls in this category. Reviews on the topic of classification can be found in many literatures (Jain, Murty et al. 1999; Jain, Duin et al. 2000).

In the statistical approach, statistical theories are used to build decision boundaries between different classes. A pattern recognition system usually contains two modes: the training mode and the testing mode. A schematic chart in Figure 5.1 shows the flow of a pattern recognition process. In the training mode, data is preprocessed; features that can represent the

sample data are extracted; and then these features are fed to a classifier for partitioning the feature space. In the testing mode, a testing data set will go through the same preprocessing and feature extraction modules and eventually be classified by the trained classifier.

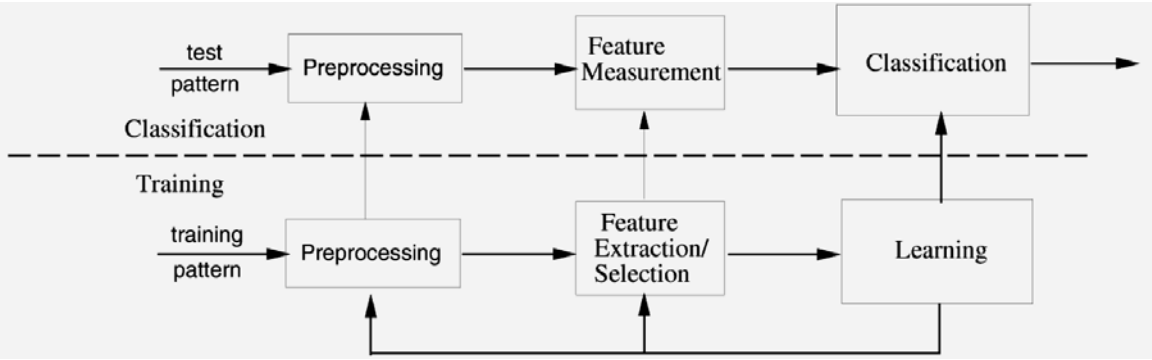

Figure 5.1 A model for statistical pattern recognition (figure is cited from (Jain, Duin et al. 2000)).

In order to evaluate the performance of a classifier, the error rate, P_e (also called the classification error), should be provided. The error rate in both training and testing data should be estimated and be taken into consideration. A number of methods for estimating the error rate exist; they differ in how the training and testing data set are split or utilized. In the n-fold cross-validation approach, the data set is divided into n subsets; (n-1) of these subsets are used as the training data and the remaining one subset is used as the testing data, and this procedure is repeated n times to give an average value of error rate. Other methods, such as leave-one-out, holdout, and rotation, can be seen as variations of n-fold cross-validation. Another useful tool to evaluate the performance of a classifier is the Receiver Operating Characteristic (ROC) Curve, which allow us to "assess the performance of the recognition system at various operating points (thresholds in the decision rule)". In a ROC plot, the True Positive Rate is plotted against the True Negative Rate. The area under the ROC curve varies between 0.0 and 1.0, while the area equals to 1.0 corresponds to 100% accuracy in classification can be achieved. In binary

classification, a random classifier would achieve a ROC of 0.5 (Zweig and Campbell 1993). In another variation of the ROC plot, the False Acceptance Rate (False Positive Rate) is plotted against the False Reject Rate (False Negative Rate) (Jain, Duin et al. 2000).

Various classification methods have been developed base on statistical theories. For feature extraction, principle component analysis (PCA), factor analysis (FA), linear discriminant analysis (LDA), multidimensional scaling (MDS), Projection Pursuit, Self-Organizing Map (SOM) and etc. are available for use. In the supervised classification approaches, classifiers can be built based on methods such as Nearest Mean Classifier, Bayes plug-in, Fisher Linear Discriminant method, Binary Decision Tree, k-Nearest Neighbor Rule, Feed Forward Neural Network, and Support Vector Machine (SVM). In the clustering approach, available methods are K-means clustering method, Single-Link or Complete-Link hierarchical clustering, Minimum Spanning Tree, Mutual Neighborhood, Mixture Decomposition, and others (Jain, Duin et al. 2000).

Support Vector Machines (SVM), or Support Vector Classifiers, was first introduced by Vapnik in mid-1990's and it became rather popular ever since then (Boser, Guyon et al. 1992; Vapnik 1995). The basic idea of SVM is to separate linearly separable patterns by finding a hyperplane with maximized margin between two patterns, where the margin is defined by data points located on the brink of each pattern and these data points are called support vectors. For non-linearly separable patterns, original data points can be transformed to a new feature space with more dimensions to achieve a separation. The transformation is implemented by a kernel function. Due to its sound theoretical foundations, performance in real-world applications, and computational efficiency, SVM has become a popular modeling and prediction method, and has been used in a variety of application areas ranging from bioinformatics (Byvatov and Schneider 2003), document classification(Joachims 1998), image retrieval (Tong and Chang 2001), and robotics (Pelossof, Miller et al. 2004).

It is worth mentioning that a data classification problem can possibly be addressed by a number of classifiers through different approaches. However, as Jain et al. had pointed out, “the choice of a classifier is a difficult problem and it is often based on which classifier(s) happen to be available, or best known, to the user” (Jain, Duin et al. 2000).

In this chapter, we describe the application of support vector machine (SVM) to classify Raman spectral data. Both Raman intensity data and intensity ratio data were included as features to represent each raw sample data. Feature selection was done through unpaired t-test and a p-value sorting procedure. Selected feature data was then fed to SVM to train a classifier, and evaluated by 10-fold cross-validation. The trained SVM classifier can predict the category of a spectrum with 85.2% accuracy. The prediction of whether the sample is from a wound that is healing or not can be 95.7% accurate when averaged spectrum of the wound is fed to the SVM classifier.

5.2 Materials and Methods

Data analysis is done in MATLAB and WEKA 3.6.2 (Waikato Environment for Knowledge Analysis, http://www.cs.waikato.ac.nz/ml/weka/) (Hall, Frank et al. 2009). All spectra collected from samples of healed patients are grouped as “healing” and all spectra collected from samples of unhealed patients “non-healing”.

Figure 5.2 shows the workflow for feature extraction and classification tasks. In the first step, starting with the pre-processed Raman spectra data set (D_0), unpaired t-tests were performed on each Raman frequency between the healing group and the non-healing group, followed by a false discovery rate (FDR) correction (Benjamini and Hochberg 1995). Data set D_0 was then sorted in ascending order according to the p-value of each frequency. Keeping only those frequencies having p-values less than 0.01 from D_0 data set yields a new data set D_1 (3037×1664). In order to remove redundancy in the feature set, in the second step, we shortened the frequency

list (1664 elements) by filtering out those frequencies that have higher p-values than their neighboring frequencies within 10 data points on the frequency list (approximately 6.8 cm^{-1}). This filtering process generated a frequency list of 101 elements. In similar Raman spectroscopy studies, it is common to seek pairwise relationships between peak intensities as discriminatory features (Utzinger, Heintzelman et al. 2001; Huang, McWilliams et al. 2003; Pichardo-Molina, Frausto-Reyes et al. 2007; Teh, Zheng et al. 2008; Crane, Brown et al. 2010). In order to discover important pairwise interactions, ratios of Raman intensities from these 101 frequencies were calculated to produce a new data set D_2 (3037×5050). This ratio data set (D_2) was also then sorted based on the p-value of t-test between the healing group and the non-healing group. The FDR correction was performed after the t-test. After removing column elements that have p-value equal or greater than 0.01, data set D_2 was updated to be a 3037 by 4350 data set. Next, D_1 and D_2 were concatenated row-wise to generate a new data set D_3. A similar processed that had been applied to data set D_1 and D_2, which included t-test, FDR correction and sorting, was applied to data set D_3 (3037×4451). In this data set, each column represents a feature of Raman spectra, which is either an intensity value of a certain frequency or a ratio of intensities between two Raman frequencies. The first 600 features in D_3 were selected to give a final feature set D_4, which was used as input data for a followed classification analysis. In the classification analysis, we used a SVM classification package provided by WEKA (weka.classifiers.functions.SMO) to classify data in D_4. This package implies Platt's Sequential Minimal Optimization (SMO) algorithm to train SVM classifier (Platt 1998; Keerthi, Shevade et al. 2001). Classifier performance was evaluated by 10-fold cross-validation.

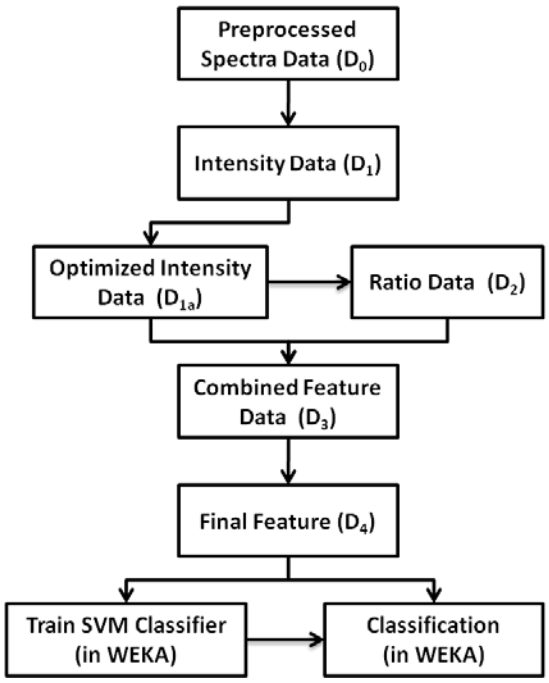

Figure 5.2 The workflow for feature extraction and classification of Raman spectral data.

5.3 Results

5.3.1 Feature Extraction and Feature Selection

In order to identify spectral features (including band intensities and intensity ratios) that can differentiate the healing wounds from the non-healing wounds, student t-tests were performed on intensity values or intensity ratio values between the two groups — healing and non-healing. Figure 5.3 compares top-5 intensity features (Figure 5.3(a)) and top-5 intensity ratio feature (Figure 5.3 (b)) between the two groups. However, note that these top-list frequencies generally do not correspond to peak positions in Raman bands, and sometimes even located in the valley of a spectrum. Frequencies showed in bar graph Figure 5.3(a) can be roughly assigned although most of these frequencies are not located on peaks: 876 and 886 cm^{-1} (C-C stretching vibration mode of hydroxyproline or tryptophan), 1276 cm^{-1} (C-N stretching and N-H bending modes of amide III of proteins), 1732 (C=O stretching vibration mode from cholesterol esters), 1420 cm^{-1}

(ring breathing or N-H in plane deformation in DNA/RNA) (Movasaghi, Rehman et al. 2007). Frequencies appearing in Figure 5.3(b) can be assigned as: 912, 924, and 960 cm^{-1} (C-C stretching in proline, hydroxyproline, or collagen backbone), 1131 cm^{-1} (C-C skeletal stretch of acyl backbone in lipids), 1394 cm^{-1} (CH_3 or CH_2 twisting in proteins and nucleic acids). The band at 1166 cm^{-1} is unassigned. A list of all single frequency features and intensity-ratio features were made by ranking the p-values from lowest to highest. The top-10 features are listed in Table 5.1. These features have low p-values due to the large number of sample size (n_1 = 1824 in the healing group; n_0 = 1213 in non-healing group). When we examine the paired ratio features in Table 5.1, feature No. 1, 4, 6, and 8 are higher value in non-healing group while the feature No. 3, 5, 7, 9, and 10 are higher in healing group.

Varying number of top significant features from data set D_3 were used to form the final feature set (D_4), which is to be used for training and testing of the support vector machine. Two different methods of selecting features were tried out.

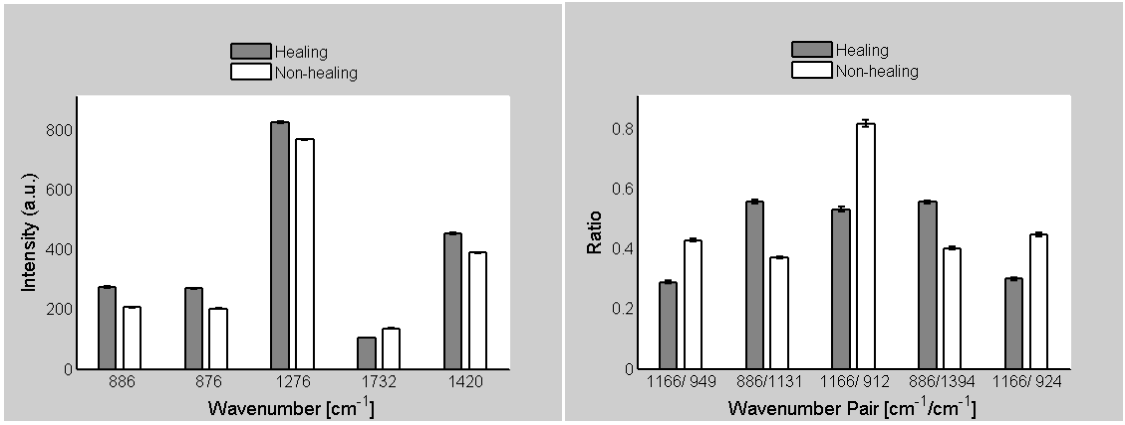

Figure 5.3 Bar graphs displaying most significant Raman band intensities (a, left) and band intensity ratios (b, right). Bars are showed as mean ± SEM. Data from healing group are showed in dark gray bars, and data from non-healing group are showed in white bars.

Table 5.1 Top 10 spectral features in data set D_3.

Ranking	Frequencies	P-value	Healing (Mean ±SEM)	Non-healing (Mean ±SEM)
1	1166/949	2.56e-099	0.29 (0.004)	0.43 (0.005)
2	886	3.39e-096	274.518 (2.158)	206.506 (2.05)
3	886/1131	4.19e-096	0.559 (0.006)	0.372 (0.004)
4	1166/912	1.49e-094	0.533 (0.008)	0.818 (0.012)
5	886/1394	7.05e-088	0.556 (0.005)	0.403 (0.005)
6	1166/924	9.41e-088	0.3 (0.004)	0.448 (0.006)
7	886/1384	2.46e-084	0.606 (0.006)	0.424 (0.005)
8	1166/960	3.47e-084	0.247 (0.003)	0.347 (0.004)
9	886/1640	3.73e-084	0.509 (0.004)	0.377 (0.004)
10	886/1371	4.84e-083	0.516 (0.006)	0.345 (0.005)

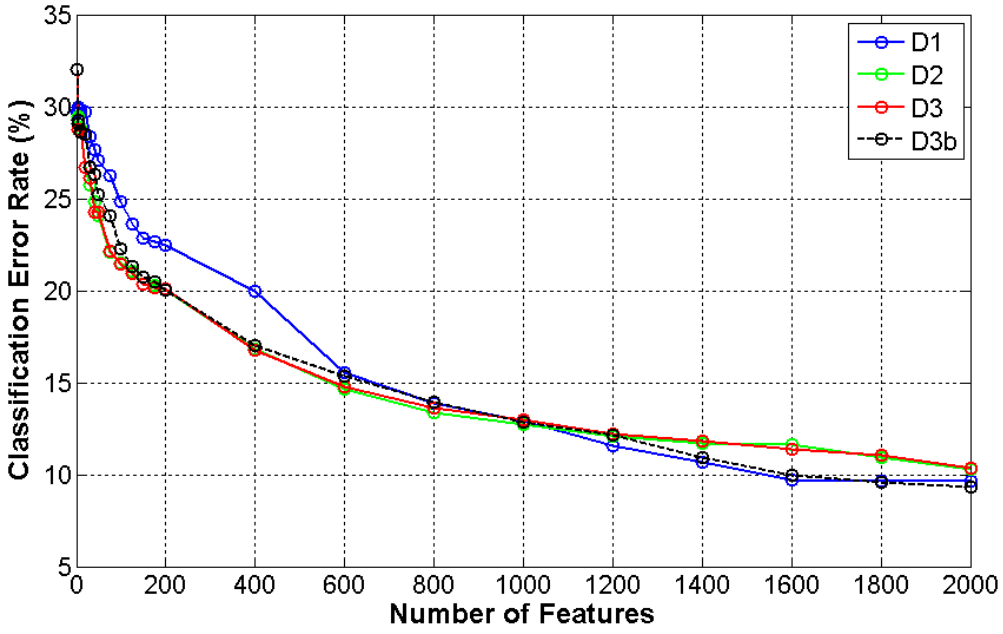

Figure 5.4 The plot for incorrectly classified instances vs. the number of features being included in data set D_4.

The first method is simply to select a specified number of features from the top of the feature-list. And for the rest of this thesis, the result is based on this method. The plot in Figure 5.4 shows a trend of incorrectly classified cases decreases as number of features being included in D_4 to train the SVM classifier. The numbers on the x-axis indicate how many features from the top of the feature list are included into final feature set (D_4). This graph shows a change in the

slope approximately at the point where the feature number is 600. Therefore 600 features from D_3 (mixed feature of D_{1a} and D_2) were selected to produce D_4 for the purpose of classification in WEKA, as described in the section of Materials and Methods. The blue and green lines in Figure 5.4 also shows similar decreasing trend in incorrectly classified cases if only features of band intensities (D_1), only features of intensity ratios (D_2), or complete single intensity feature mixed with intensity ratio features (D_{3b}, features are also ordered by p-values) are included in D_4. Using features from D_2 for classification gives results that are almost as good as using features from D_3. Classification error of data set D_1, D_2, and D_{3b} are shown in Figure 5.4 only for comparison purpose, while only D_3 is used for the final feature set D_4.

In the second method, each feature in the final feature set (D_4) is picked from the feature data set (D_3), evaluated by the classification result from the SVM classifiers, and then added to the final feature set (D_4). More specifically, we start off with a D_4 containing only one feature, which gives the best SVM classification results of all the features in D_3. Next, we pick another feature from D_3 as the second feature in D_4, and use SVM classification result as the criteria to decide which feature from D_3 should be kept as the second feature in D_4. This selection process is repeated to pick more features from D_3. In all the SVM classification procedures, a 10-fold cross-validation method is used. Due to the fact that D_3 contains 4351 features, it's computationally expensive to test all features in D_3. Therefore, feature selection is limited to only the top 1000 features in data set D_3 and 50 features are picked to form final feature set D_4. The first 10 features selected by this method are shown in Table 5.2, in which the first column gives the sequence number of a feature in D_4 and the second column gives the ranking of the feature in data set D_3. In contrast to Table 5.1 where the top-10 features used for method 1 are listed, features shown in Table 5.2 are mostly not high ranked features. The Figure 5.5 shows a plot of classification error versus the number of features (from 1 to 50) included in D_4 using both feature selecting methods. The second method always has smaller classification error rate at given number of features in D_4,

therefore with a less amount of features selected by method 2 we can achieve a classification accuracy that otherwise requires more features in data set D_4. It's also notable that the curve for method 2 reaches a shorter plateau region on the curve from x = 7 to x = 14. Similarly, there is another plateau where the classification error rate remains at about 21%, starting from x = 17 until the end of curve. However, we know from the first method (Figure 5.4) that the classification error rate can be decreased to about 15% when the first 600 features from D_3 are used in D_4. But the line for method 2 doesn't fit in a simple linear, quadratic, nor exponential line model, and therefore it's difficult to conjecture when can the value of classification error rate drop to about 15%. A prominent drawback of applying the second selecting method is that it is computationally expensive; it takes long time to select its first 50 features from the first 1000 features in D_3 data set. Therefore, in this study, we choose to use the first feature selection method and continue our analysis based on the first method.

Table 5.2 List of the first 17 features selected by the second method.

#	Ranking	Frequencies	P-value	Healing (Mean ±SEM)	Non-healing (Mean ±SEM)
1	9	886/1640	3.73e-84	0.509 (0.004)	0.377 (0.004)
2	504	1732/701	5.04e-45	0.827 (0.012)	1.139 (0.02)
3	703	949/668	5.92e-40	2.071 (0.019)	1.709 (0.017)
4	719	876/997	9.74e-40	0.928 (0.02)	0.583 (0.012)
5	954	855	2.38e-35	589.31 (5.063)	495.629 (5.445)
6	213	642/1394	3.78e-56	0.569 (0.007)	0.404 (0.006)
7	150	886/971	1.23e-59	0.693 (0.009)	0.479 (0.008)
8	783	642/1229	2.03e-38	0.409 (0.007)	0.282 (0.005)
9	287	1582/924	1.97e-52	1.653 (0.018)	2.076 (0.02)
10	449	886/997	8.37e-47	0.908 (0.017)	0.582 (0.011)

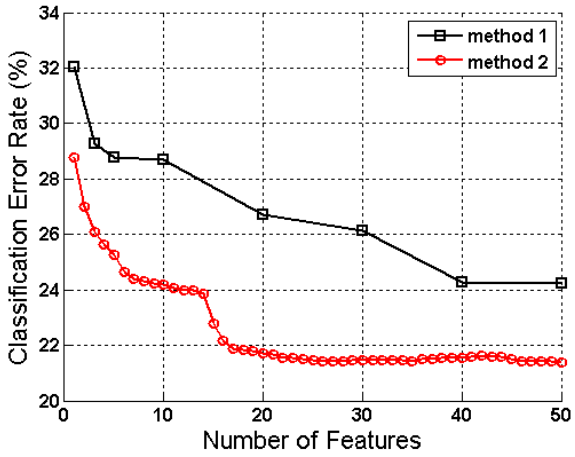

Figure 5.5 The change of classification error against the change of the number of features contained in the final feature set (D_4).

5.3.2 Spectral Data Classification

Thus, the final data set D_4 is formed by selecting the first 600 features from D_3 data set. Classification result is evaluated by 10-fold cross-validation and the result is shown in Table 5.3. The total correctly classified the instances are 2588, accounts for 85.2% of total spectra. To further examine the prediction result on samples, we show the error rate for each sample in Figure 5.6. The average error rates of samples in each group are showed in Figure 5.6 (a). The Incorrectly Classified Cases (Error Rate) of each sample are ordered in time and are plotted in Figure 5.6 (b). While the left plot shows the error rate of all healing and non-healing samples, the right plot shows error rates of each sample only from the healing group. Data points in both plots are order by time from the earliest time of sample collecting to the last time of sample collecting. The classification errors of healing group are generally lower than that of the non-healing group. However, there is no apparent time-related trend that is being showed in both plots.

Table 5.3 Confusion matrix of the classification result for all spectra.

		Prediction		Summary
		Healing	**Non-healing**	
Actual Group	**Healing**	1656 (91%)	168 (9%)	1824
	Non-healing	281 (23%)	932 (77%)	1213
Summary		1937	1100	3037

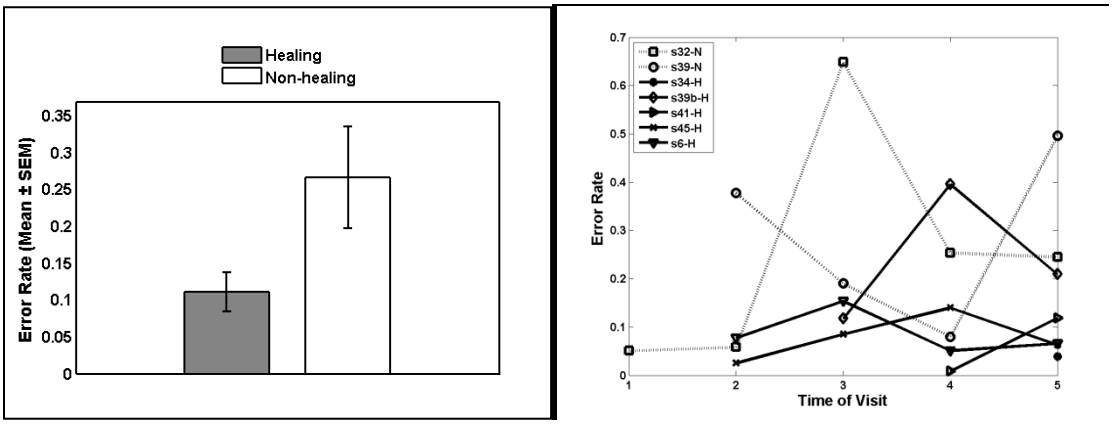

Figure 5.6 The classification error rate of spectra from each sample. (a) averaged classification error rate (mean±SEM); (b) classification error rate of spectra from each sample organized by the subject and ordered by the sequence that each sample is being collected.

Although we have the predicted category for each individual spectrum, we are more interested in predicting the category for each sample. One way to achieve this goal is to make a decision about the sample's category based on predictions for all the spectra collected from this sample. In our computing procedure, the category is represented numerically, where "healing" is represented as "1" and "non-healing" is represented as "0". So when we take an averaged of predicted numerical category values for all spectra from a sample, we can get a score whose value falls between 0 and 1. Then we have to decide the sample's category using this score. A straight forward way is to round this score so we get either "0" or "1" as predicted category for this sample. This step is equivalent to threshold the score at 0.5. To find out how the thresholding

process could affect the overall prediction results to these samples, we tried a series of threshold values (from 0 to 1, step size is 0.05) and show the performance in a receiver operating characteristic (ROC) plot (Figure 5.7). As showed in Figure 5.7, the integrated area under the curve is 0.948, suggesting that category of a sample can be predicted from its Raman spectra with high accuracy. More specifically, when the threshold is between 0.50 and 0.75 only one sample is misclassified, corresponding to an accuracy of 95.7%. When the threshold is between 0.5 and 0.6, one non-healing sample is predicted as healing, and when the threshold is between 0.65 and 0.75, one healing sample is predicted as non-healing.

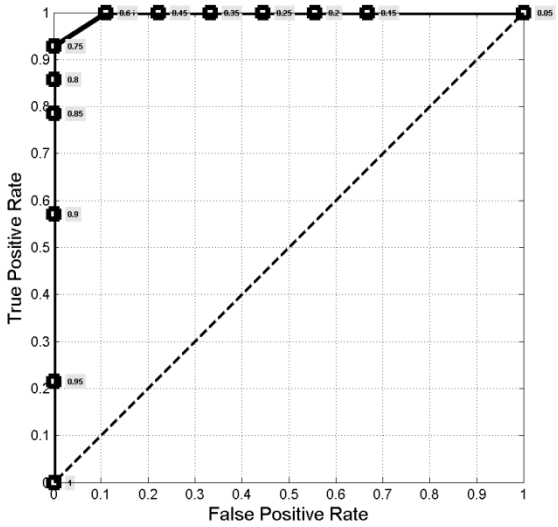

Figure 5.7 The receiver operating characteristic (ROC) curve of changing threshold values. From the top-right corner moving toward the bottom-left corner, the threshold values labeled on the data points are 0.05, 0.2 (0.1, 0.15), 0.25 (0.2), 0.35 (0.3), 0.45 (0.4), 0.6 (0.5, 0.55), 0.75 (0.65, 0.7), 0.8, 0.85, 0.9, and 0.95. Threshold values in the parenthesis give the same true positive rate/false positive rate ratio.

The alternative way to predict the category of a sample is to take an average of all individual data from the sample and then classify the averaged data with the trained SVM classifier. In this way, only one non-healing sample was misclassified as healing wound sample,

giving us an accuracy of classification of 95.7% (Table 5.4). This classification result is virtually the same as the result we get from the previous thresholding method when the threshold is between 0.5 and 0.6. In other tentative classification trials, we included the first 800 or more features from data set D_3 to comprise D_4. All of those trials give us 100% accuracy in classifying 23 samples and the area under the ROC curve is 1(the ROC curve is not shown here). The Classification Error Rate for 23 samples when different number of feature is included in the final data set (D_4) is shown in Figure 5.8. As shown in the plot, when 800 or 1000 features are included, all 23 samples can be correctly classified.

Table 5.4 Confusion matrix of classification results for averaged spectra from each sample when 600 features are included in D_4.

		Prediction		**Summary**
		Healing	**Non-healing**	
Actual Group	**Healing**	14	0	14
	Non-healing	1	8	9
		15	8	23

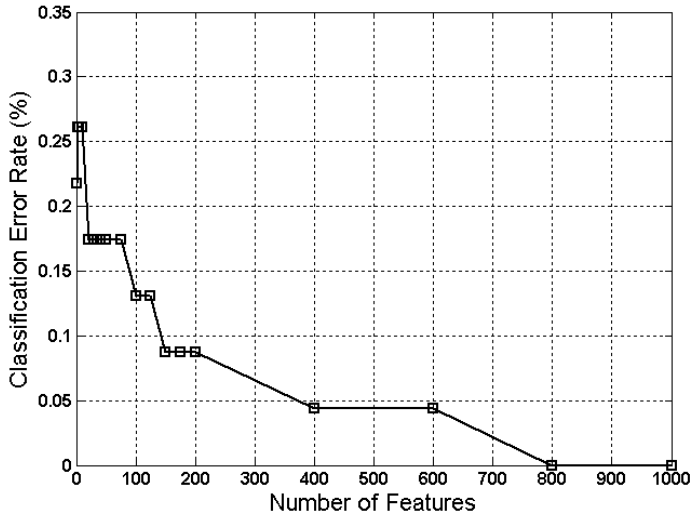

Figure 5.8 The plot of incorrectly classified samples vs. the number of features being included in the final data set (D_4) for training the SVM classifier. The trained classifier was then used to classify 23 samples. The number of features (x-axis) ranges from 1 to 1000.

5.4 Wound Color Analysis

In addition to the spectral data analysis and classification, we are also interested to find out if the color of wounds can help differentiate healing wounds from non-healing wounds. So we analyzed wound color from 23 images of wounds from which we collected wound debridement samples. These images were collected by the same method and under the same condition as previously have been described in the chapter 3 (Chapter 3.2 Materials and Methods). The wound boundary in each image was manually traced to segment wound area from the rest of image. The mean numerical values of colors in 7 color channels (R, G, B, L*, a*, b*, ΔE) were then calculated for the wound area (Note that in this section, ΔE is the Euclidean color distance of a color relative to a RGB reference color (0, 0, 0), which is different from the ΔE in the chapter 3). The image processing was done in MATLAB (version 7.14 (R2012a)). Wound images were labeled as healing wound or non-healing wound based on the actual wound label. T-tests were performed between the two groups on each color channel to evaluate whether there was any

significant color difference. A trial SVM classification of these 23 wound images was also attempted using wound color values in these 7 color channels.

The p-values are shown in Table 5.5, where none of these color channels show significant difference between the healing and the non-healing group. The classification result of SVM is shown in the Table 5.6. As seen in the confusion table, all cases were predicted as healing, which resulted in a 39.1% classification error rate. This result is only slightly better than a random guess, and the classifier wasn't separating these two groups. Both t-test and SVM classification results from the current 23 wound images imply that the color of wound bed is probably not a good differentiate factor for healing and non-healing wounds.

Table 5.5 P-values of t-test in color channels between wound images of healing and non-healing group.

	R	**G**	**B**	**L***	**a***	**b***	**ΔE**
P-values	0.4559	0.2730	0.4921	0.3185	0.8746	0.4833	0.4937

Table 5.6 The classification result of wounds using color information of the wound area.

		Prediction		**Summary**
		Healing	**Non-healing**	
Actual Group	**Healing**	14	0	14
	Non-healing	9	0	9
		23	0	23

5.5 Discussion

Raman spectra analysis was usually done on peak intensity or intensity ratio. Single peak intensity gives information of the content of certain molecular structure relative to the entire Raman active substance, while the intensity ratio gives information about the relative content of the two molecular structures contributing to the two peaks. The use of intensity ratio (I_{1445}/I_{1655}) as a diagnostic marker has been reported by Huang et al. (Huang, McWilliams et al. 2003) and by Crane et al. (I_{1665}/I_{1445}) (Crane, Brown et al. 2010). Utzinger et al. also used two pairs of intensity ratios (I_{1330}/I_{1454}, and I_{1454}/I_{1656}) in their spectra analysis as diagnostic markers for detection of cervical precancers (Utzinger, Heintzelman et al. 2001). Teh et al. also mentioned utilizing I_{875}/I_{1450} ratio as a marker for classification (Teh, Zheng et al. 2008). Pichardo-Molina et al. compared a series intensity ratio values between cancer and control group (Pichardo-Molina, Frausto-Reyes et al. 2007). In this study, we include both peak intensities and peak intensity ratios into the feature list, on which the Support Vector Machine was trained.

Among these top-list features, which are ranked by p-values, intensity-ratios are dominant. This can be explained by the notion that ratios can reflect the relative content of a certain substance in the sample and therefore shows a more consistent differences between groups, while the intensity of a single frequency can be affected by the overall signal level in one spectrum and are less likely to show consistent differences.

As we can see from the Table 5.1, frequency 886 cm^{-1} appears as the top single peak intensity feature and appears as one of the peak in many intensity-ratio pairs. Since it's identified as the most differentiable single intensity feature between groups and it's actually a point in the valley, we conjecture that it is a reasonable and reliable reference peak to compare intensities of other frequencies. Among those frequencies paired with 886 cm^{-1} in the list, only 1131 and 1394 are peaks, other frequencies such as 1371, 1384, and 1640 cm^{-1} are located on slopes of adjacent

bands. Note that at locations of these non-peaked frequencies, changes on the gradient of slope can be seen, a sign of weak band overlapped with strong band. The band at 1131 cm^{-1} is assigned to C-N stretching in proteins, or C-C skeletal stretch associated with acryl backbone transconformation (Movasaghi, Rehman et al. 2007). The intensities from healing group at 1371, 1384, and 1394 cm^{-1} are all higher than that of non-healing group, however it is difficult to decide molecular assignments since multiple vibration modes, such as ring breathing mode of DNA/RNA bases, C-H rocking, and C-N stretching, can contribute to their signal. At frequency of 1640 cm^{-1}, both amide I band and intermolecular bending mode of water can contribute to the signal. Due to the indefinite assignment of these bands, no conjecture of pathological difference between healing and non-healing groups can be made.

Also can be seen in Table 5.1 is that four pairs of intensity ratio contains the frequency of 1166 cm^{-1}, which is located in a valley. The other four frequencies that pairs with it ranges from 912 to 960 cm^{-1}, which corresponds to a band peaking at 938 cm^{-1} (912-952 cm^{-1}). The band at 938 cm^{-1} can be assigned to C-C stretching mode of proline, hydroxyproline or collagen backbone (Movasaghi, Rehman et al. 2007). The intensities in this spectra region from the healing group are all higher than that of the non-healing group while the intensity at 1166 cm^{-1} in healing group are lower than non-healing group, which explains why ratio values from these pairs are low in healing group comparing to non-healing group (see Figure 5.3(b)). Therefore, these ratio features can be associated with a conjecture of high collagen content in healing group.

After examining the list (Table 5.1), we think the sorting process using p-values did pick out features that are reasonable in terms of Raman spectra and biochemical composition. However, these features are still quite difficult to be clearly interpreted in the sense of tissue's biochemical composition.

Quite a few studies on Raman spectroscopy on biological samples aimed at finding a method to separate one group from another, or even identifying diagnostic markers. The complex and heterogeneous nature of biological tissue requires a decent amount of Raman spectra to be collected. Meanwhile, it also drives people seeking solutions from multivariate analysis and data mining methods. Currently, most the classification analysis on Raman spectra didn't include intensity ratio data. Methods such as k-means clustering, LDA, SVM, and decision trees are largely applied to intensity data or dimension reduced data by PCA. But as potent diagnostic markers, band intensity ratios should be included into the process of multivariate analysis or data classification. We show that including intensity ratios as features improves the classification accuracy.

In classification, there are always concerns about over-training the classifier and long computation time. Given that there are 1767 frequencies recorded in between the range of 604 - 1798cm^{-1}, it would be impractical and unnecessary to calculate intensity ratios for all possible frequency combinations. So we shortened the list down to 101 frequencies by selecting most significant but non-adjacent frequencies, which greatly reduced the number of combinations of frequency pair. The data size is further reduced by selecting only the first 600 most significant features, intensity or intensity ratio, based on the t-test results. The high percentage of intensity ratios on the 600-feature list practically proves that intensity ratio is important spectral marker. Also note that when using the same number of features in the training data for classifier, classification outcome is much better if intensity ratio data is included in the data set (Figure 5.4). On the contrary, whether to include single intensity data in the training set will not make a big difference, as we can see from lines for D_2 and D_3 in Figure 5.4, since they only account for a small percentage of features in the top-600 feature list. However, we also note that there is no clear cut on how many features should be choose according to the plot in Figure 5.4, except a

slight change in the slope of line for D_3 at the point of 600-features. As more features are included in training data set D_4, the classification outcomes are improved.

When the error rate of classification outcome is organized by sample collection sequence, no clear trend can be seen (Figure 5.6(b)). The classification error of samples from the last clinical visit is not necessary smaller than samples from earlier clinical visits. This might be due to the fluctuation of healing status of wounds, the debridement samples were not a constant good representation of wounds, or because the classifier is classifying spectra in a way that matches the true pathological changes happening in wound bed.

As for the classification results for 23 averaged spectra, the only one misclassified sample was a non-healing sample but predicted as healing sample. This error might root in the inconsistent quality of debridement samples. While most of debridement samples are small pieces of solid tissue, this sample appeared to be a non-solid mixture scraped from wound surface. The cryosectioned slides of this sample appeared to be nonconsecutive tissue fragments rather than an integral piece under the microscope and only small amount of spectra could be collected. This sample is an unfaithful representation of the wound and eventually leads to a classification error. However, note that even this sample can be correctly classified as the number of feature in training data was increased.

In summary, we used a SVM classifier in WEKA (SMO algorithm). We introduced the methods of using intensity ratio values together with spectra intensity to train SVM classifier. Classification result is evidently improved by the induction of intensity ratio data. The trained SVM classifier can predict the category of a spectrum with 85.2% accuracy. The prediction of the category of a sample is of 95.7% accuracy when the averaged spectrum of each wound sample is fed to the trained SVM classifier. In section 5.4, possible causes of misclassification are analyzed.

Only one sample out of 23 was not correctly classified. However, the classification accuracy can reach 100% as the number of features used in training data increased.

Chapter 6: Conclusions and Future Work

6.1 Conclusions

The work in this thesis focused on improving evaluation of wound healing status by applying current engineering techniques, such as image processing, near infrared Raman spectroscopy, and data classification.

The image processing method developed for detecting wound is a semi-automatic method which requires only simple manual operations, yet it can effectively detect wounds on wound images from patients with various skin colors. Our method uses thresholding for grayscale image from individual color channels as well as pixel-color-comparison for the color image to generate binary wound images, and then combines them with predetermined channel weights to produce a final binary wound image, from which wound boundary and wound size can be determined. We evaluated the performance of this method by applying it to detect wound boundaries in wound images of animal models and human subjects with chronic wounds, and then comparing the computer detected wound boundaries with manually traced boundaries. Our evaluation results show that this method performs very well on our animal wound images. The performance is slightly compromised on human wound images due to the heterogeneity of the color on the wound bed as well as on the surrounding skin.

Our work on Raman spectroscopy for wound samples, as described in chapter 4, has demonstrated a method to study chronic wounds by Raman spectroscopy. The protocol of preparing wound samples ensures a good sampling of wound samples with Raman spectra. Although Raman spectra from both healing and non-healing wounds samples share similar spectral profiles, the average spectra from these two groups show some differences. While the

average spectrum of healing samples has relatively higher intensities at bands associated with proteins, especially collagen, average spectra of non-healing samples is relatively high in intensities at bands associated with cell nucleus.

In Chapter 5, a classification method is developed to classify wound samples by their Raman spectra. Statistical analyses of spectral intensity data show that even though average spectra of healing and non-healing samples have significant difference at most frequencies between 600-1800 cm^{-1}, the two groups are not separable at any single frequency due to big variations within each group. Similar statistical result is also found in the intensity ratio data. A supervised classification method, support vector machine (SVM), is applied to selected features including both intensities and intensity ratios. The top-600 features, including both intensities and intensity ratios, are extracted from the original spectral data to represent each spectral measurement and then used for classification. The SVM classifier is trained and evaluated by 10-fold cross-validation. Of all 3037 measurements, 85.2% can be correctly classified. When the average spectra of each sample are used for classifying the sample, 22 out of 23 samples are correctly classified (95.7% accuracy).

6.2 Future Work

Future directions for wound image processing include: (a) application of machine learning methods to generate the binary wound images, and optimization of the weight for each binary image in order to improve the accuracy of the final wound boundary; (b) further reduction of the manual operations during the process; (c) development of a graphical user interface (GUI) to increase the ease of operations; (d) extraction of the wound texture information as well as color information for further characterization or classification.

On Raman spectroscopy study of wounds, further work is needed to confirm the results we have observed in the current study. A larger patient population would provide stronger

statistical power for the distinguishing spectral features of non-healing wounds. The consistency of the wound sample can be improved by enforcing that only samples with consistent good quality will be collected and used for the study. Further histological evidence is needed to confirm the biochemical differences suggested by the Raman spectral analysis. Lastly, other Raman Spectra analysis methods can be applied, such as the Band Target Entropy Minimization (BTEM) method developed to extract information about individual molecular components from the Raman spectral data (Widjaja and Garland 2010).

An integrated approach to wound healing assessment may involve incorporation of data collected from multiple modalities. Current technologies enable us to collect information on wound size, color, shape, and visual patterns through image analysis, on molecular constituents through Raman spectroscopy, and on tissue oxygenation information through diffuse near infrared spectroscopy (DNIRS) (Neidrauer 2010). Even more information about the patient can be retrieved from the medical records and other demographic data. With a comprehensive set of information, a thorough characterization of the wound and a more reliable assessment regarding the healing status and treatment options would be possible.

List of References

Acha, B., C. Serrano, et al. (2005). "Segmentation and classification of burn images by color and texture information." Journal of Biomedical Optics **10**: 034014.

Alberti, K. G. and P. Z. Zimmet (1998). "Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation." Diabet Med **15(**7): 539-553.

Apelqvist, J., K. Bakker, et al. (2008). "Practical guidelines on the management and prevention of the diabetic foot: based upon the International Consensus on the Diabetic Foot (2007) Prepared by the International Working Group on the Diabetic Foot." Diabetes Metab Res Rev **24 Suppl 1**: S181-187.

Askanazi, J., P. Furst, et al. (1980). "Muscle and plasma amino acids after injury: hypocaloric glucose vs. amino acid infusion." Ann Surg **191**(4): 465-472.

Ball, D. W. (2001). "Rayleigh and Raman scattering." Spectroscopy **16**(2): 28-+.

Baum, C. L. and C. J. Arpey (2005). "Normal cutaneous wound healing: clinical correlation with cellular and molecular events." Dermatol Surg **31**(6): 674-686; discussion 686.

Belem, B. (2004). Non-invasive wound assessment by image analysis Ph.D. Thesis, University of Glamorgan, UK.

Beljebbar, A., O. Bouche, et al. (2009). "Identification of Raman spectroscopic markers for the characterization of normal and adenocarcinomatous colonic tissues." Critical Reviews in Oncology Hematology **72**(3): 255-264.

Benjamini, Y. and Y. Hochberg (1995). "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." Journal of the Royal Statistical Society. Series B (Methodological) **57**(1): 289-300.

Benson, H. A. E. and A. C. Watkinson (2011). Topical and transdermal drug delivery : principles and practice. Hoboken, N.J., Wiley.

Boser, B. E., I. M. Guyon, et al. (1992). A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory. Pittsburgh, Pennsylvania, United States, ACM**:** 144-152.

Boulton, A. J. M., L. Vileikyte, et al. (2005). "The global burden of diabetic foot disease." The Lancet **366**(9498): 1719-1724.

Brem, H., P. Sheehan, et al. (2004). "Protocol for treatment of diabetic foot ulcers." Am J Surg **187**(5A): 1S-10S.

Byvatov, E. and G. Schneider (2003). "Support vector machine applications in bioinformatics." Applied bioinformatics **2**(2): 67-77.

Calhoun, J. H., K. A. Overgaard, et al. (2002). "Diabetic foot ulcers and infections: current concepts." Adv Skin Wound Care **15**(1): 31-42; quiz 44-35.

CDC (2011). National diabetes fact sheet: national estimates and general information on diabetes and prediabetes in the United States, 2011. U. S. D. o. H. a. H. S. Centers for Disease Control and Prevention. Atlanta, GA, Centers for Disease Control and Prevention.

Chan, J. W., D. S. Taylor, et al. (2006). "Micro-Raman spectroscopy detects individual neoplastic and normal hematopoietic cells." Biophys J **90**(2): 648-656.

Chatfield, C. and A. J. Collins (1980). Introduction to multivariate analysis. London ; New York, Chapman and Hall.

Cheng, W. T., M. T. Liu, et al. (2005). "Micro-Raman spectroscopy used to identify and grade human skin pilomatrixoma." Microsc Res Tech **68**(2): 75-79.

Crane, N. J., T. S. Brown, et al. (2010). "Monitoring the healing of combat wounds using Raman spectroscopic mapping." Wound Repair Regen **18**(4): 409-416.

Crow, P., A. Molckovsky, et al. (2005). "Assessment of fiberoptic near-infrared Raman spectroscopy for diagnosis of bladder and prostate cancer." Urology **65**(6): 1126-1130.

Deng, J. L., Q. Wei, et al. (2005). "Study of the effect of alcohol on single human red blood cells using near-infrared laser tweezers Raman spectroscopy." Journal of Raman Spectroscopy **36**(3): 257-261.

Deshpande, M. M. (2008). "Diabetic foot ulcer: assessment and management." J Indian Med Assoc **106**(6): 400.

Duckworth, M., N. Patel, et al. (2007). A clinically affordable non-contact wound measurement device. Proceedings of 30th RESNA conference on technology and disability, , Phoenix, USA, .

Ellis, D. I. and R. Goodacre (2006). "Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy." Analyst **131**(8): 875-885.

Enoch, S. and P. Price (2004). "Cellular, molecular and biochemical differences in the pathophysiology of healing between acute wounds, chronic wounds and wounds in the aged." World Wide Wounds.

Falanga, V. (2005). "Wound healing and its impairment in the diabetic foot." Lancet **366**(9498): 1736-1743.

Faolain, E. O., M. B. Hunter, et al. (2005). "Raman spectroscopic evaluation of efficacy of current paraffin wax section dewaxing agents." J Histochem Cytochem **53**(1): 121-129.

Filik, J. and N. Stone (2008). "Analysis of human tear fluid by Raman spectroscopy." Analytica Chimica Acta **616**(2): 177-184.

Fisher, T. K., R. Wolcott, et al. (2010). "Diabetic foot infections: A need for innovative assessments." Int J Low Extrem Wounds **9**(1): 31-36.

Gantwerker, E. A. and D. B. Hom (2011). "Skin: histology and physiology of wound healing." Facial Plast Surg Clin North Am **19**(3): 441-453.

George, B., M. Cebioglu, et al. (2010). "Inadequate diabetic care: global figures cry for preventive measures and personalized treatment." The EPMA Journal **1**(1): 13-18.

Gill, S. E. and W. C. Parks (2008). "Metalloproteinases and their inhibitors: regulators of wound healing." Int J Biochem Cell Biol **40**(6-7): 1334-1347.

Gniadecka, M., P. A. Philipsen, et al. (2004). "Melanoma diagnosis by Raman spectroscopy and neural networks: Structure alterations in proteins and lipids in intact cancer tissue." Journal of Investigative Dermatology **122**(2): 443-449.

Gurtner, G. C., S. Werner, et al. (2008). "Wound repair and regeneration." Nature **453**(7193): 314-321.

Hall, M., E. Frank, et al. (2009). "The WEKA data mining software: an update." SIGKDD Explor. Newsl. **11**(1): 10-18.

Han, H. W., X. L. Yan, et al. (2009). "Analysis of serum from type II diabetes mellitus and diabetic complication using surface-enhanced Raman spectra (SERS)." Applied Physics B-Lasers and Optics **94**(4): 667-672.

Hanlon, E. B., R. Manoharan, et al. (2000). "Prospects for in vivo Raman spectroscopy." Phys Med Biol **45**(2): R1-59.

Harding, K. G., H. L. Morris, et al. (2002). "Science, medicine and the future: healing chronic wounds." BMJ **324**(7330): 160-163.

Hirobe, T. (1983). "Proliferation of epidermal melanocytes during the healing of skin wounds in newborn mice." Journal of Experimental Zoology **227**(3): 423-431.

Huang, W. E., R. I. Griffiths, et al. (2004). "Raman microscopic analysis of single microbial cells." Anal Chem **76**(15): 4452-4458.

Huang, Z. W., A. McWilliams, et al. (2003). "Near-infrared Raman spectroscopy for optical diagnosis of lung cancer." International Journal of Cancer **107**(6): 1047-1052.

Ikoma, T., H. Kobayashi, et al. (2003). "Physical properties of type I collagen extracted from fish scales of Pagrus major and Oreochromis niloticas." Int J Biol Macromol **32**(3-5): 199-204.

Jain, A. K., R. P. W. Duin, et al. (2000). "Statistical pattern recognition: A review." Ieee Transactions on Pattern Analysis and Machine Intelligence **22**(1): 4-37.

Jain, A. K., M. N. Murty, et al. (1999). "Data clustering: A review." Acm Computing Surveys **31**(3): 264-323.

Jeffcoate, W. J. and K. G. Harding (2003). "Diabetic foot ulcers." Lancet **361**(9368): 1545-1551.

Jessup, R. L. (2006). "What is the Best Method for Assessing the Rate of Wound Healing?: A Comparison of 3 Mathematical Formulas." Advances in Skin Wound Care **19**(3): 138-147.

Joachims, T. (1998). Text categorization with Support Vector Machines: Learning with many relevant features. Machine Learning: ECML-98. C. Nédellec and C. Rouveirol, Springer Berlin / Heidelberg. **1398:** 137-142.

Kantor, J. and D. J. Margolis (2000). "A multicentre study of percentage change in venous leg ulcer area as a prognostic index of healing at 24 weeks." British Journal of Dermatology **142**(5): 960-964.

Keast, D. H., C. K. Bowering, et al. (2004). "MEASURE: A proposed assessment framework for developing best practice recommendations for wound assessment." Wound Repair and Regeneration **12**(s1): 1-17.

Keerthi, S. S., S. K. Shevade, et al. (2001). "Improvements to Platt's SMO Algorithm for SVM Classifier Design." Neural Computation **13**(3): 637-649.

Khavkin, J. and D. A. F. Ellis (2011). "Aging Skin: Histology, Physiology, and Pathology." Facial Plastic Surgery Clinics of North America **19**(2): 229-234.

Kolesnik, M. and A. Fexa (2004). Segmentation of wounds in the combined color-texture feature space. SPIE, Medical Imaging.

Kolesnik, M. and A. Fexa (2005). Multi-dimensional Color Histograms for Segmentation of Wounds in Images. Image Analysis and Recognition**:** 1014-1022.

Kolesnik, M. and A. Fexa (2006). How robust is the SVM wound segmentation? Proceedings of the 7th Nordic Signal Processing Symposium, 2006**:** 50-53.

Koljenovic, S., L. P. Choo-Smith, et al. (2002). "Discriminating vital tumor from necrotic tissue in human glioblastoma tissue samples by Raman spectroscopy." Laboratory Investigation **82**(10): 1265-1277.

Krafft, C., T. Knetschke, et al. (2006). "Studies on stress-induced changes at the subcellular level by Raman microspectroscopic mapping." Anal Chem **78**(13): 4424-4429.

Krafft, C., T. Knetschke, et al. (2003). "Mapping of single cells by near infrared Raman microspectroscopy." Vibrational Spectroscopy **32**(1): 75-83.

Ladwig, G. P., M. C. Robson, et al. (2002). "Ratios of activated matrix metalloproteinase-9 to tissue inhibitor of matrix metalloproteinase-1 in wound fluids are inversely correlated with healing of pressure ulcers." Wound Repair Regen **10**(1): 26-37.

Lait, M. E. and L. N. Smith (1998). "Wound management: a literature review." Journal of Clinical Nursing **7**(1): 11-17.

Lazarus, G. S., D. M. Cooper, et al. (1994). "Definitions and guidelines for assessment of wounds and evaluation of healing." Wound Repair Regen **2**(3): 165-170.

Leung, P. C. (2007). "Diabetic foot ulcers--a comprehensive review." Surgeon **5**(4): 219-231.

Lieber, C. A. and A. Mahadevan-Jansen (2003). "Automated method for subtraction of fluorescence from biological Raman spectra." Applied Spectroscopy **57**(11): 1363-1367.

Liu, Y., D. Min, et al. (2009). "Increased matrix metalloproteinase-9 predicts poor wound healing in diabetic foot ulcers: Response to Muller et al." Diabetes Care **32**(11): e137.

Lobmann, R., A. Ambrosch, et al. (2002). "Expression of matrix-metalloproteinases and their inhibitors in the wounds of diabetic and non-diabetic patients." Diabetologia **45**(7): 1011-1016.

Manoharan, R., J. J. Baraga, et al. (1992). "Quantitative histochemical analysis of human artery using Raman spectroscopy." J Photochem Photobiol B **16**(2): 211-233.

Margolis, D. J., L. Allen-Taylor, et al. (2002). "Diabetic Neuropathic Foot Ulcers: the association of wound size, wound duration, and wound grade on healing." Diabetes Care **25**(10): 1835-1839.

Margolis, D. J., J. M. Gelfand, et al. (2003). "Surrogate End Points for the Treatment of Diabetic Neuropathic Foot Ulcers." Diabetes Care **26**(6): 1696-1700.

Martin, P. (1997). "Wound healing--aiming for perfect skin regeneration." Science **276**(5309): 75-81.

McCreery, R. L. (2000). Raman spectroscopy for chemical analysis. New York, John Wiley & Sons.

Mehendale, F. and P. Martin (2001). The cellular and molecular events of wound healing. Cutaneous Wound Healing. V. Falanga. London, United Kingdom, Taylor & Francis.

Mekkes, J. R. and W. Westerhof (1995). "Image processing in the study of wound healing." Clinics in Dermatology **13**(4): 401-407.

Movasaghi, Z., S. Rehman, et al. (2007). "Raman spectroscopy of biological tissues." Applied Spectroscopy Reviews **42**(5): 493-541.

Muller, C. A., O. Belyaev, et al. (2007). "Corticosteroid-binding globulin: a possible early predictor of infection in acute necrotizing pancreatitis." Scand J Gastroenterol **42**(11): 1354-1361.

Neidrauer, M. (2010). Assessment of Chronic Wounds Using In Vivo Diffuse Near Infrared Spectroscopy. Ph.D., Drexel University.

Ohta, N. and A. R. Robertson (2005). Colorimetry: Fundamentals and Applications. Chichester, England ; Hoboken, N.J. , John Wiley & Sons Ltd.

Omberg, K. M., J. C. Osborn, et al. (2002). "Raman spectroscopy and factor analysis of tumorigenic and non-tumorigenic cells." Applied Spectroscopy **56**(7): 813-819.

Ortiz, C., D. Zhang, et al. (2006). "Validation of the drop coating deposition Raman method for protein analysis." Anal Biochem **353**(2): 157-166.

Papazoglou, E. S., M. S. Weingarten, et al. (2008). Assessment of diabetic foot ulcers with diffuse near infrared methodology. BioInformatics and BioEngineering, 2008. BIBE 2008. 8th IEEE International Conference on.

Parks, W. C. (1999). "Matrix metalloproteinases in repair." Wound Repair Regen **7**(6): 423-432.

Pelossof, R., A. Miller, et al. (2004). An SVM learning approach to robotic grasping. Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on.

Perez, A. A., A. Gonzaga, et al. (2001). Segmentation and analysis of leg ulcers color images. International Workshop on Medical Imaging and Augmented Reality 262-266.

Pichardo-Molina, J. L., C. Frausto-Reyes, et al. (2007). "Raman spectroscopy and multivariate analysis of serum samples from breast cancer patients." Lasers in Medical Science **22**(4): 229-236.

Plassmann, P. and T. D. Jones (1998). "MAVIS: a non-invasive instrument to measure area and volume of wounds." Medical Engineering & Physics **20**(5): 332-338.

Plataniotis, K. N. and A. N. Venetsanopoulos (2000). Color Image Processing and Applications. Berlin; Heidelberg, Springer.

Platt, J. (1998). Fast Training of Support Vector Machines using Sequential Minimal Optimization. Advances in Kernel Methods - Support Vector Learning. B. Schoelkopf, C. Burges and A. Smola, MIT Press.

Quan, S. Y., G. S. Lazarus, et al. (2007). "Digital Imaging of Wounds: Are Measurements Reproducible Among Observers?" International Journal of Lower Extremity Wounds **6**(4): 245-248.

Ramsey, S. D., K. Newton, et al. (1999). "Incidence, outcomes, and cost of foot ulcers in patients with diabetes." Diabetes Care **22**(3): 382-387.

Reeds, P. J., C. R. Fjeld, et al. (1994). "Do the differences between the amino acid compositions of acute-phase and muscle proteins have a bearing on nitrogen loss in traumatic states?" J Nutr **124**(6): 906-910.

Reiber, G. E., B. A. Lipsky, et al. (1998). "The burden of diabetic foot ulcers." Am J Surg **176**(2A Suppl): 5S-10S.

Robson, M. C. and A. Barbul (2006). "Guidelines for the best care of chronic wounds." Wound Repair and Regeneration **14**(6): 647-648.

Robson, M. C., D. P. Hill, et al. (2000). "Wound Healing Trajectories as Predictors of Effectiveness of Therapeutic Agents." Archives of Surgery **135**(7): 773-777.

Rook, A. and T. Burns (2004). Rook's textbook of dermatology. Malden, Mass., Blackwell Science.

Russ, J. C. (2002). The image processing handbook. Boca Raton, FL, CRC Press.

Schanda, J. (2007). CIE colorimetry. Colorimetry: Understanding the CIE System. J. Schanda. New Jersey, John Wiley & Sons Inc.

Schultz, G. S., R. G. Sibbald, et al. (2003). "Wound bed preparation: a systematic approach to wound management." Wound Repair and Regeneration **11**(s1): S1-S28.

Schuster, K. C., I. Reese, et al. (2000). "Multidimensional information on the chemical composition of single bacterial cells by confocal Raman microspectroscopy." Anal Chem **72**(22): 5529-5534.

Sheehan, P. (2006). "Percent Change in Wound Area of Diabetic Foot Ulcers over a 4-Week Period Is a Robust Predictor of Complete Healing in a 12-Week Prospective Trial." Plastic and Reconstructive Surgery **117**(7S): 239S-244S.

Shetty, G., C. Kendall, et al. (2006). "Raman spectroscopy: elucidation of biochemical changes in carcinogenesis of oesophagus." British Journal of Cancer **94**(10): 1460-1464.

Shim, M. G. and B. C. Wilson (1997). "Development of an in vivo Raman spectroscopic system for diagnostic applications." Journal of Raman Spectroscopy **28**(2-3): 131-142.

Sikirzhytski, V., K. Virkler, et al. (2010). "Discriminant Analysis of Raman Spectra for Body Fluid Identification for Forensic Purposes." Sensors **10**(4): 2869-2884.

Singer, A. J. and R. A. Clark (1999). "Cutaneous wound healing." N Engl J Med **341**(10): 738-746.

Singh, N., D. G. Armstrong, et al. (2005). "Preventing foot ulcers in patients with diabetes." JAMA **293**(2): 217-228.

Snyder, R. G., H. L. Strauss, et al. (1982). "Carbon-hydrogen stretching modes and the structure of n-alkyl chains. 1. Long, disordered chains." The Journal of Physical Chemistry **86**(26): 5145-5150.

Sowemimo, G. O. A., J. Naim, et al. (1982). "Repigmentation after burn injury in the guinea-pig." Burns **8**(5): 345-357.

Stadelmann, W. K., A. G. Digenis, et al. (1998). "Physiology and healing dynamics of chronic cutaneous wounds." Am J Surg **176**(2A Suppl): 26S-38S.

Steed, D. L., C. Attinger, et al. (2006). "Guidelines for the treatment of diabetic ulcers." Wound Repair and Regeneration **14**(6): 680-692.

Teh, S. K., W. Zhene, et al. (2010). "Near-infrared Raman spectroscopy for optical diagnosis in the stomach: identification of Helicobacter-pylori infection and intestinal metaplasia." International Journal of Cancer **126**(8): 1920-1927.

Teh, S. K., W. Zheng, et al. (2008). "Diagnostic potential of near-infrared Raman spectroscopy in the stomach: differentiating dysplasia from normal tissue." British Journal of Cancer **98**(2): 457-465.

Teh, S. K., W. Zheng, et al. (2010). "Near-infrared Raman spectroscopy for early diagnosis and typing of adenocarcinoma in the stomach." British Journal of Surgery **97**(4): 550-557.

Thawer, H. A., P. E. Houghton, et al. (2002). "A comparison of computer-assisted and manual wound size measurement." Ostomy/Wound Management **48**(10): 46-53.

Tong, S. and E. Chang (2001). Support vector machine active learning for image retrieval. Proceedings of the ninth ACM international conference on Multimedia. Ottawa, Canada, ACM**:** 107-118.

Trengove, N. J., M. C. Stacey, et al. (1999). "Analysis of the acute and chronic wound environments: the role of proteases and their inhibitors." Wound Repair and Regeneration **7**(6): 442-452.

Utzinger, U., D. L. Heintzelman, et al. (2001). "Near-infrared Raman spectroscopy for in vivo detection of cervical precancers." Applied Spectroscopy **55**(8): 955-959.

Vapnik, V. N. (1995). The nature of statistical learning theory. New York, Springer.

Velander, P., C. Theopold, et al. (2008). "Impaired wound healing in an acute diabetic pig model and the effects of local hyperglycemia." Wound Repair Regen **16**(2): 288-293.

Weingarten, M. S., M. Neidrauer, et al. (2010). "Prediction of wound healing in human diabetic foot ulcers by diffuse near-infrared spectroscopy: a pilot study." Wound Repair Regen **18**(2): 180-185.

Weingarten, M. S., E. Papazoglou, et al. (2006). "Measurement of optical properties to quantify healing of chronic diabetic wounds." Wound Repair and Regeneration **14**(3): 364-370.

Weingarten, M. S., E. S. Papazoglou, et al. (2008). "Correlation of near infrared absorption and diffuse reflectance spectroscopy scattering with tissue neovascularization and collagen concentration in a diabetic rat wound healing model." Wound Repair and Regeneration **16**(2): 234-242.

Whitney, J., L. Phillips, et al. (2006). "Guidelines for the treatment of pressure ulcers." Wound Repair and Regeneration **14**(6): 663-679.

Widgerow, A. D. (2011). "Chronic wound fluid--thinking outside the box." Wound Repair Regen **19**(3): 287-291.

Widjaja, E. and M. Garland (2010). "Detection of bio-constituents in complex biological tissue using Raman microscopy. Application to human nail clippings." Talanta **80**(5): 1665-1671.

Widjaja, E., G. H. Lim, et al. (2008). "A novel method for human gender classification using Raman spectroscopy of fingernail clippings." Analyst **133**(4): 493-498.

Wild, T., M. Prinz, et al. (2008). "Digital measurement and analysis of wounds based on colour segmentation." European Surgery **40**(1): 5-10.

Wood, B. R., B. Tait, et al. (2001). "Micro-Raman characterisation of the R to T state transition of haemoglobin within a single living erythrocyte." Biochimica Et Biophysica Acta-Molecular Cell Research **1539**(1-2): 58-70.

Zhang, L., M. J. Henson, et al. (2005). "Multivariate data analysis for Raman imaging of a model pharmaceutical tablet." Analytica Chimica Acta **545**(2): 262-278.

Zweig, M. H. and G. Campbell (1993). "Receiver Operating Characteristic (Roc) Plots - a Fundamental Evaluation Tool in Clinical Medicine (Vol 39, Pg 561, 1993)." Clinical Chemistry **39**(8): 1589-1589.

Vita

Xiang Mao

Education

- **Ph.D.** Biomedical Engineering Jun. 2012
  Drexel University, PA
- **M.S.** Polymer Chemistry and Physics Apr. 2005
  Beijing University of Aeronautics and Astronautics (Beihang University), China
- **B.S.** Materials Science and Engineering Jul. 2002
  Beijing University of Aeronautics and Astronautics (Beihang University), China

Publications and Conferences

- **Mao, X.**, Sacan, A., et al., *Supervised Classification of Raman Spectra from Diabetic Foot Ulcers*. (In preparation).
- Papazoglou, E.S., Zubkov, L., **Mao, X**., et al., *Image analysis of chronic wounds for determining the surface area.* Wound Repair and Regeneration, 2010. **18**(4): p. 349-358. (as the main contributor)
- Weingarten, M.S., Neidrauer, M., Mateo, A., **Mao, X.**, et al., *Prediction of wound healing in human diabetic foot ulcers by diffuse near-infrared spectroscopy: A pilot study*. Wound Repair and Regeneration, 2010. **18**(2): p. 180-185.
- **Mao, X.**, Papazoglou, E.S., "Monte Carlo Simulation for Chronic Wound Assessment with Diffuse NIR Spectroscopy", BMES Annual Conference (Austin, TX) (2010).
- He, J., Wu, Y.Z., Wu, J., **Mao, X.**, et al., *Study and Application of a Linear Frequency−Thickness Relation for Surface-Initiated Atom Transfer Radical Polymerization in a Quartz Crystal Microbalance.* Macromolecules, 2007. **40**(9): p. 3090-3096.
- **Mao, X.**, Huang P.C., *Preparation and Characterization of the Inclusion Complex of p-Diethynylbenzene with β-Cyclodextrin*, Chinese Journal of Applied Chemistry, 2005. **22**(10): p. 1080-1082.
- **Mao, X.**, Huang P.C., Wang W.Z., Zhang W., *Synthesis and the fluorescence properties of a new oligo(p-phenyleneethynylene)*, Journal of Functional Polymer, 2005.