Design and Evaluation of a Visible-to-Near-infrared Spectrograph for Grain Quality Assessment

BY

Wenbo Wang

A Thesis submitted to
the Faculty of Graduate Studies
In Partial Fulfillment of the Requirements
for the Degree of

MASTER OF SCIENCE

Department of Biosystems Engineering
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

Design and Evaluation of a Visible-to-Near-infrared Spectrograph for Grain Quality Assessment

BY

Wenbo Wang

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Master of Science

ABSTRACT

Grain grading using electronic tools based on near-infrared spectroscopy and machine vision has the potential to replace current visual inspection methods. A compact and robust lab-scale spectrograph system that can be customized for agricultural applications was designed and tested. The spectrograph can detect spectral signals with wavelengths ranging from approximately 433 nm to 1067 nm. Spectral data in the short near-infrared region (866 - 1067 nm) and in the visible region (433 - 750 nm) can be acquired simultaneously. The spectrograph was configured to work under reflectance mode with external illumination and a full spectral scan required 0.7 s. Data acquisition and user interface programs were developed under LabVIEW environment.

A 650 nm single-mode diode laser was used to calibrate the spectrograph wavelength equation. An argon lamp was used to validate calibrated wavelength equation. A maximum wavelength error of 1.95 nm was determined. An optical bandpass expressed in full-width-at-half-maximum (FWHM) of 5.7 nm in the 433 - 750 nm, and 5.5 nm in the 866 - 1067 nm region was estimated experimentally. Repeated measurement of reflectance spectra of a Polytetrafluoroethylene (PTFE) disk provided coefficients of variation (CV) values below 0.2% within 30 s sampling period and remained under 0.4% within 30 min sampling period. To test the applicability of the spectral signals colleted using the spectrograph, moisture content of ground wheat samples was determined in the 866 nm to 1067 nm region. For an independent validation set, root mean squared error of prediction (RMSEP) of 1.3% and coefficient of determination (R^2) around 0.88 were achieved.

ACKNOWLEDGEMENTS

Having a chance to undertake graduate study in Canada is a dream for many people like me. Therefore, I would like to take this opportunity to thank my advisor, Dr. Jitendra Paliwal, for helping me fulfill this academic dream. The vision, support, encouragement, and great understanding that Dr. Paliwal has given me during my two years of study were instrumental to the successful completion of this program. I could never thank him enough.

I am thankful to Dr. D.S. Jayas and Dr. T. Beta for serving on my advisory committee. I also wish to express my gratitude towards Dr. L. Leonardi at the Institute for Biodiagnostics (IBD), National Research Council (NRC), who brought in his equipment to assist with system testing. I am also thankful to Dr. F. Jian at the Department of Biosystems Engineering, who gave me helpful suggestions on the preparation of wheat samples.

I thank Mr. Dale Bourns and Mr. Matt McDonald for their technical assistance. They demonstrated their professional skills and offered me lots of practical ideas throughout the project, which was quite a learning experience as well.

The project was supported by the Faculty of Agricultural and Food Sciences and the University Research Grants Program of the University of Manitoba. I would also like to acknowledge the partial funding provided by the Canada Research Chairs (CRC) program and Natural Sciences and Engineering Research Council of Canada (NSERC). The Canadian Wheat Board (CWB) is acknowledged for infrastructure support.

I would also like to take this chance to thank my friends both in Canada and China, who have shown me the emotional support and sincere friendship that inspired me along the way.

Finally, I would like to thank my beloved parents, to whom I owe everything. It took them so much sacrifice to make me a better person that I could not pay them back by any means.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Grain Handling and Grading in Canada: Present and Future

As one of the leading grain producing countries in the world, Canada has a total grain and oilseeds (including winter wheat, spring wheat, durum wheat, oats, barley, rye, flaxseed, canola, corn, and soybeans) production of 57.8 million tonnes on average annually (Canada Grains Council 2003). Approximately, an average of 25.2 million tonnes of these grain and oilseeds are exported annually through a grain collection, handling, and distribution system (Canada Grains Council 2003). To ensure the quality, the grains to be exported have to undergo strict quality inspection administered by the Canadian Grain Commission (CGC). Nowadays, grains from farms are collected and transported using trucks to primary elevators, where these grains are sorted by types before being transferred to terminal elevators using railcars (Luo 1997). At the terminal elevators, trained inspectors from CGC perform the classification and grading of grains using visual inspection methods. Some final quality control operations are then carried out before they are exported (Lásztity and Salgó 2002).

The grain production and its export is an important industry to Canada, and grain quality is the deterministic factor for export success of Canadian grain. To remain competitive in the international cereal grain market, Canada must take measures to ensure the quality and at the same time increase the grain handling throughput capability. Till now, the rigorously regulated quality assurance system by the federal government has earned Canada a worldwide reputation in providing millions of tons of grains and oilseeds with consistent end-use quality from year to year. At the turn of 21st century,

Canada's grading and handling system anticipates the changes necessary to accommodate the need in the future marketing of grain, which tends to direct purchase on specifications, or even by variety, instead of grade. The functional and analytical aspects of grain are becoming more important in grain marketing rather than the visual characteristics (Williams 1999).

Although, establishment of the relationship of visual grades to actual functionality has been attempted by researchers, there is still no conclusive result yet. However, most grain with different end-use characteristics is currently predicted solely based on visual inspection methods. In Canada, elevator managers and CGC inspectors distinguish the class based on a combination of seed-coat color and physical kernel shapes and sizes that are distinctive. Consequently, a grade for the grain shipment is established based on a set of grade standards, namely, test weight, varietal purity, soundness, vitreousness, and maximum limit of foreign material (excluding dockage). Under the present system of visual grading, the most important parameters affecting appearance of the grain are fairly loosely defined. Also, due to factors such as personal judgment, work conditions, level of fatigue, etc., consistent results among individual grain inspectors are not always easy to achieve. In today's grain marketing, cereal processors need to know the functionality of grain that can comply with the requirements of their clients. To provide consistent, objective, and accurate information on the grain to be marketed, it is desirable to have an electronic grading system capable of determining the visual as well as physico-chemical properties of cereal grains.

1.2 Recent Research Progress and Potential Research Activities

Recent progress in the development of electronics, optics, and computing technologies has enabled machine-vision to move its steps into various industrial fields such as automotive, manufacturing, and electronics. The last decade has witnessed tremendous progress in applying machine-vision techniques to the agri-food industry in Canada. Previous research activities placed an emphasis on image analysis algorithm development, morphological feature extraction, and pattern recognition using statistical methods and artificial neural networks (ANNs) (Luo 1997; Paliwal 2002; Visen 2002).

Although near-infrared spectroscopy (NIRS) technique has already established itself in many different agri-food fields to analyze components such as protein, fat, starch, moisture, etc. (Osborne 2000), further research is still required to implement the technique into the process of grain grading.

The future of an electronic grain grading system will rely on combination of machine-vision and NIRS techniques, which are both technically mature to a considerable degree, independently. The combined imaging technique, also referred to as hyperspectral imaging technique, can provide necessary information on functionality and analytical aspects of cereal grain products, simultaneously. Currently, major obstacles in implementing hyperspectral imaging technique for automated grain grading are prohibitive equipment cost and lack of applied research.

1.3 Near-infrared Spectroscopy as a Viable Technology for Grain Quality Assessment

The reason that NIRS technique is so widely accepted in the grain industry is that it possesses technical merits that satisfy many quality inspection needs in a modern grain

handling and grading system. The main advantages that make NIRS such an attractive analytical technique are as follows:

1. Results of NIRS technique are objective, reliable, and consistent. The reason is that near-infrared (NIR) spectroscopic instruments are highly automated, stable, and spectra measurement is highly reproducible. Moreover, modern NIRS instruments agree among each other to a high degree.
2. Near-infrared Spectroscopy technique is fast and non-destructive. Recording of spectra is less time consuming and there is less need for sample preparation. For example, seeds in breeding program do not have to be destroyed to determine its fertility. Moreover, the NIR spectra contain a wealth of information to determine multiple physico-chemical properties simultaneously and thus greatly reduce the complexity of analysis task.
3. Since it is basically an optical method, NIRS doesn't pose any environmental hazard compared to conventional methods that involve chemical processes and reagents.
4. Instrument functions can be easily upgraded. Little hardware modification is required during upgrade and instrument function can be expanded by simply sharing mathematical calibration models for different physico-chemical properties or different grain products.
5. As an instrumental technique for routine analysis tasks, NIRS saves labor and reduces the cost. Little training is required for instrument operators and there is minimal need of human intervention once the whole system has been calibrated.

No chemical consumable is involved in NIR analysis and the samples, generally, do not have to be destroyed.

Despite its wide acceptance in grain-based industry and technical advantages that fit into the needs of an electronic grading system, research needs to be done before it can be integrated with machine-vision technique to replace current visual grading system and be practically implemented in on-site grain quality inspection system in Canada.

1.4 Objectives of Research

Application of the machine-vision technique in grain quality assessment has been worked on for many years and research outcomes are at the verge of actual industrial application. Gaining better appreciation of NIRS technique can benefit the development of a suitable electronic grain grading system in the future. The focus of this research is to construct and evaluate a spectrograph system capable of working in the visible-to-NIR region. A spectrograph configuration is chosen because of its low construction cost, portability, rugged structure, and flexibility in performance. These characteristics are expected to meet the demands of future generations of NIRS equipments in the agricultural industry.

The applicability of using NIRS to quantitatively determine the chemical composition of wheat and its many other physical or chemical properties had already been proven by a list of publications as reviewed by McClure (2003). Thus, a preliminary experiment was conducted to determine the moisture level in ground wheat using their reflectance spectra in an effort to demonstrate the potential use of the spectrograph system for grain quality assessment. The main focus of the thesis, however, is to design and test the spectrograph rather than moisture determination of wheat using NIRS.

The specific research objectives were:

1. to design the optical system and mechanical mounting devices, provide theoretical calculations, and construct a diffraction grating spectrograph;
2. to configure the data acquisition system, and develop spectral data logging and monitoring program using LabVIEW;
3. to calibrate wavelength with regards to CCD array detector output and validate the wavelength relationship, and to determine the system spectral width experimentally;
4. to carry out a preliminary experiment using the constructed NIR spectrograph to acquire spectra of ground whole wheat, compare two reflectance spectra preprocessing methods, and develop linear calibration models to quantitatively determine the moisture content of ground wheat.

2. LITERATURE REVIEW

2.1 History of Near-infrared Spectroscopy

The NIR region extends from 780 nm to 2500 nm in wavelength (or from 12,800 – 4000 cm^{-1} in wavenumber). This spectral region was first discovered by Fredrick William Herschel in year 1800 (Herschel 1800a). In an experiment that Sir Herschel ran to find out the heating efficiency of different colors of light dispersed from white sunlight, he discovered that temperature continued to rise beyond the red color region with a maximum heating effect into the black area (Herschel 1800b). He named the region beyond the red color as "calorific rays" and sometime later after his discovery, this invisible region was referred to as infrared (Pasquini 2003). The discovery of infrared radiation intrigued the interest of scientists to study the electromagnetic radiation and its interaction with substances. In year 1850, Bunsen and Kirchhoff built the first spectrometers working in the visible region. From then on, optical spectroscopy has evolved into a distinct scientific activity.

Despite the popularity of NIRS today, spectroscopic study in the mid-infrared region extending from 2500 nm to 25,000 nm first gained wide acceptance. Year 1889 marked the beginning of infrared spectroscopy when the first spectra within the mid-infrared range were recorded (Ogilvie 1989). In 1900, Coblentz obtained absorbance spectra of pure substances and applied them in identification of organic functional groups. Due to the huge number of unknown chemical compounds, a large library of spectra had to be set up and a match was found whenever a spectrum of unknown compound was recorded and needed to be identified. Mid-infrared spectroscopy thus had been an important tool in identification of unknown chemicals (Stewart 1970). More details on

the history of infrared absorption spectroscopy, on discovery of absorbance spectra of pure substances, and instrumentation development in the early years of spectroscopy have been compiled by Jones (1985).

For the first few decades in the early 20th century, mid-infrared spectroscopy continued to receive constant research and instrumentation development. On the other hand, study of the light absorption in the NIR region was largely neglected by researchers in optical spectroscopy. The fact is that NIR spectra are complex, unintelligible, and occupied by broad, highly overlapping, and weak absorption bands (Miller 2001). This had practically diverted the interest of most practitioners of spectroscopy, and thus had hindered the research progress to find practical applications of NIRS. Until 1950s, only a few NIR spectroscopic systems were investigated and the application of NIRS did not begin until the computer was invented (Davies 1998). In year 1983, professor David Wetzel suggested that the great potential of NIRS had long been neglected describing NIRS as the "sleeper among spectroscopic techniques" (Wetzel 1983). In the last quarter of the 20th century, researchers began to realize the importance of NIRS technique, which was evidenced by a boom in the number of publications on NIRS in the 1980s. The total number of papers published dealing with NIRS was about 255 from year 1930 to 1980, and this number nearly quadrupled as the result of intensified research activities in NIRS (Burns and Ciurczak 1992). Owing to the fast development in electronics, photonics, and computer technologies, inexpensive and robust NIRS systems could be made readily available to industry. Meanwhile, development and application of multivariate mathematical techniques known as chemometrics made the revolution of

NIRS technique possible. In the 1990s, NIRS technique had become so widely exposed that it found practical applications in virtually all possible industries.

Workman (1999) documented the utilization of NIRS techniques in fields like agriculture (Johnsen 1997; Vazquez de Aldana et al. 1996), animal Sciences (Carter et al. 1996), biotechnology (Berger et al. 1996; Macaloney et al. 1997), biomedical (Zhang et al. 2002), earth sciences, atmospheric science and mineralogy (Malley et al. 1996; Matuszewska 1996), environmental science (Eilert et al. 1995; Somsen et al. 1996), fine chemicals and chemical production (Ge et al. 1995; Wang et al. 1995), food and beverages (Hewavitharana and Van Brakel 1997; McQueen et al. 1995), forensic science (Dal Cason 1997), instrument development physics (Gallaher et al. 1996; Shaffer and Small 1996), medical and clinical chemistry (Brennan et al. 1997; Franck et al. 1996), petroleum and fuel research (Workman 1996), pharmaceutical production (Blanco et al. 1996), plant sciences, pulp and paper (Sun et al. 1997), polymer science (Ozpozan and SchraderKeller 1997), semiconductors and materials science (Sperline et al. 1995), and textiles (Kitagawa et al. 1997).

2.2 Theory of Near-infrared Spectroscopy

The most important aspect of NIRS as an analytical tool is that it can determine the chemical composition and physicochemical behavior of foods and their raw materials. This is due to the fact that NIRS analyzes the sample in a way that reflects the actual number of molecules of individual constituents in the sample (Murray and Williams 1990). It is known that all organic matter consists of atoms, mainly carbon, oxygen, hydrogen, nitrogen, phosphorus, and sulphur, with minor amounts of other elements. These atoms combine by covalent and electrovalent bonds to form molecules (Campbell

et al. 2002). Without external radiation, the molecules vibrate in their fundamental energy levels at ambient temperature. When radiated using a light source with continuous spectral output, only light at particular wavelengths is absorbed. The energy of photons at those wavelengths corresponds to the energy gaps between two fundamental energy levels, or overtone, and combinations of vibration levels. Absorption of light in the NIR region involves transfer of radiation energy into mechanical energy associated with the motion of atoms bonded together by chemical bonds, and this process can be explained using theory of vibrational spectroscopy, the classical mechanical model, and the anharmonic model.

2.2.1 Vibrational spectroscopy

Photons can be absorbed by molecules, and the absorption of photons results in the change of vibrational or rotational energy increase of the molecules. According to Einstein, the energy of a photon at a given wavelength (λ) is given by:

$$E = h\nu \tag{2.1}$$

where E is the energy of a single photon (J), ν is the frequency of the light in vacuum (Hz), and h is Planck's constant (6.63×10^{-34} J·s). It is also observed that the absorption of photons by molecules is selective, that is, the molecules can not take up photon energy continuously. Molecules can only absorb photon packets that are of an energy level coincident with the characteristic vibrations of the molecule and result in the molecules being excited to a higher energy level.

2.2.2 The harmonic model and anharmonic model

Hooke's law has been used to determine the vibration frequency of the "ball and spring" model (Raghavachari 2001). It is a simple and straightforward way to look at the molecular vibration using such a mechanical model that describes the diatomic molecule as two masses attached to ends of a spring with a force constant (k). Any pair of bound atoms in a molecule exhibits a natural oscillation and vibrates at a specific frequency given by:

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \tag{2.2}$$

where k is the force constant ($N \cdot m^{-1}$), and μ (kg) is the reduced mass given by:

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \tag{2.3}$$

where m_1 and m_2 are two masses in a diatomic molecule (kg).

The potential energy of the system depends upon the square of the displacement from equilibrium and is continuous. A diatomic molecule model substitutes the atoms for balls and chemical bond for the spring and can be used to explain some spectral phenomena. However, unlike the mechanical spring model, the fact is that matter can not absorb energy continuously and the quantum-mechanical model is introduced. In 1900, Max Planck proposed quantum theory stating that the energy levels of such an oscillator are quantized at equally spaced values and the transitions can only occur between two adjacent levels (Osborne et al. 1993). The discrete energy level (E_v) is given by Eq. 2.4:

$$E_v = (\upsilon + \frac{1}{2})h\nu \tag{2.4}$$

where υ is the vibrational quantum number (υ = 0, 1, 2, ...) and E_υ is the energy associated with the υth quantum level (J). Energy transition between two adjacent energy levels requires absorption of electromagnetic radiation whose energy is identical to the difference between two adjacent energy levels given by:

$$\Delta E = \Delta\upsilon \cdot h\nu = h\nu \tag{2.5}$$

where ΔE is the energy difference (J), $\Delta\upsilon$ is the quantum number difference and $\Delta\upsilon$ equals ±1 in the quantum-mechanical model. However, because the quantum/harmonic model forbids transition with $\Delta\upsilon$ greater than 1, the overtone bands in the NIR region that have $\Delta\upsilon$ equal to 2 or greater can not be explained. Moreover, according to the theory, two or more vibrations can not interact with each other to produce combination bands either.

An empirical function known as the Morse Function, is used to approximate the anharmonic behavior of a diatomic molecule using the equation:

$$E = E_D(1 - e^{-\alpha y})^2 \tag{2.6}$$

where E_D is the dissociation energy (J), α is a constant for a particular molecule (m^{-1}), and y is the displacement of atoms from equilibrium position (m). Using Eq. 2.6 to solve the mechanical equations, the solution for the anharmonic oscillator becomes

$$E = (\upsilon + 0.5)h\nu - (\upsilon + 0.5)^2 h\nu x - (\upsilon + 0.5)^3 h\nu x' - \ldots \tag{2.7}$$

where x, x'...are know as anharmonicity constants of vibration with decreasing values between 0.005 and 0.05. For small values of quantum number υ, the third and beyond terms in Eq. 2.7 are ignored, and therefore, we have:

$$E \approx (\upsilon + 0.5)h\nu - (\upsilon + 0.5)^2 h\nu x \tag{2.8}$$

The anharmonic/quantum model can be used to explain the occurrence of transitions with $\Delta v \geq 2$ and existence of combination bands. These two types of bands are the most common absorption bands in the NIR spectral region. Curves depicting energy of a diatomic molecule undergoing simple harmonic and anharmonic vibration are shown in Fig. 2.1. More details regarding the theory of vibrational NIR spectroscopy can be found in other references (Workman 1996; Bokobza 1998).

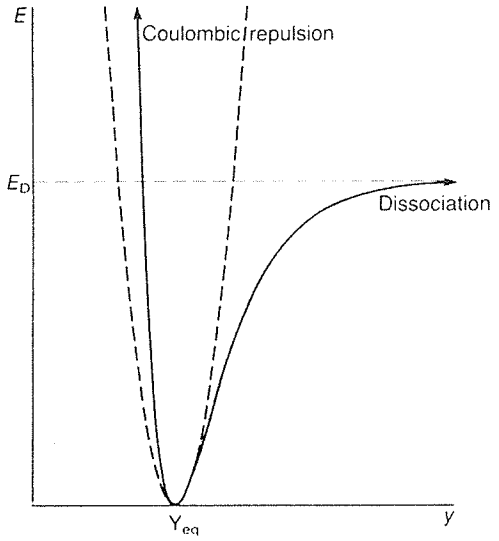


Figure 2.1 The energy of a diatomic molecule undergoing simple harmonic motion (dotted curve) and anharmonic vibration (solid curve). *Source:* Osborne et al. (1993)

2.3 Interaction of Radiation with Matter and Measurement Modes

Various substances exist in different forms and have different response under infrared radiation. When electromagnetic radiation interacts with a sample, it may be absorbed, transmitted or reflected. Transparent substances transmit and absorb light easily. For opaque materials, a large portion of light is reflected or diffuse reflected due to scattering. Based on sample properties and forms of propagation of NIR light in the sample, there exist several measurement modes as depicted in Fig. 2.2.

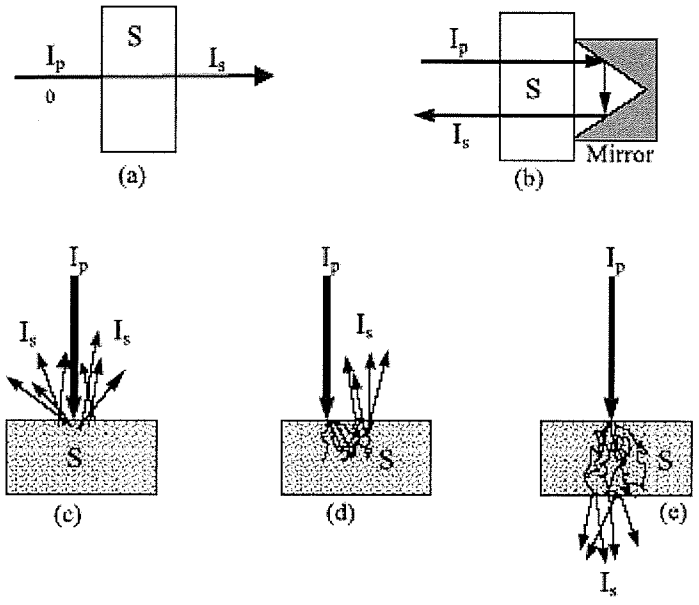


Figure 2.2 Modes of measurement employed in NIR spectroscopy a) transmittance; b) transflectance; c) diffuse reflectance; d) interactance; and e) diffuse transmittance through scattering medium. *Source:* Pasquini (2003)

Transmittance is applied to measure transparent samples possessing minimum light scattering effect. Usually, sample in liquid form or solvent is presented in glass or quartz cell since NIR light is transparent to glass. The fraction of transmitted radiation (I_s/I_p) by the sample is called transmittance. In practice, transmittance is converted to absorbance given by the following relationship:

$$A = \log(\frac{1}{T}) = \log(\frac{I_p}{I_s}) \tag{2.9}$$

Where A is the absorbance in absorbance units (AU), T is the transmittance (no unit), I_s is the incoming light energy (J), and I_p is the transmitted light energy (J).

The relation of concentration of a sample, the sample thickness, and the absorbance is governed by Beer-Lambert law (Swinehart 1972):

$$A = abc \tag{2.10}$$

where a is a constant called the absorptivity ($L \cdot mol^{-1} \cdot m^{-1}$), b is the sample thickness (m), and c is the concentration of a sample (mol/L).

Transflectance is a modified version of transmittance. A retro-reflector is often employed behind the sample cuvette to double the optical path length through the sample.

Diffuse reflectance is applied to measurement of solid samples and is perhaps the most accepted measurement mode in NIR spectroscopy. Kubelka-Munk function has been introduced to describe the energy of reflected radiation using two constants called the scattering constant (s) and the absorption constant (k). For a special case of an opaque layer of infinite thickness, the relation could be given by the equation (Kubelka and Munk 1931):

$$F(R_\infty) = \frac{(1-R_\infty)^2}{2R_\infty} = \frac{k}{s} = \frac{ac}{s} \tag{2.11}$$

where $F(R_\infty)$ is the Kubelka-Munk function (no unit), R_∞ is the reflectance of the infinitely thick layer (no unit), k is the absorption constant (mm^{-1}), and s is the scattering constant (mm^{-1}) (Birth and Zachariah 1976).

Apparent absorbance as given in Eq. 2.12 is used in practice instead of Kubleka-Munk function.

$$A_R = \log(\frac{1}{R_\infty}) \tag{2.12}$$

where A_R is apparent absorbance also in absorbance units (AU) and is assumed to be proportional to concentration (c).

Interactance is a modified version of diffuse reflectance. The collected radiation signal travels a much longer distance in the sample and is assumed to be richer in information on sample constituent than that collected under diffuse reflectance mode.

Diffuse transmittance is different from diffuse reflectance in that diffuse transmittance signal is collected after light travels through the sample and emerges on the other side of sample. This mode is often used at short NIR wavelengths with turbid liquid or solid sample with a thickness of 10 - 20 mm.

2.4 Instrumentation

2.4.1 Overview of spectroscopic instrument

Practical application of NIR spectroscopy has been around for several decades, and there has been a wide array of instruments available for different end-user purposes. Each kind of instrument is based on different working principles and possesses certain performance characteristics. Currently in the agricultural and food related fields, spectroscopic instruments can be put into two categories:

Dispersive systems: One type of dispersive system is based on diffraction gratings. According to its instrument configuration, this type of spectrometer can be divided into scanning monochromator and spectrograph. Scanning monochromator works by mechanically rotating the diffraction grating to tune the wavelength of light to be received by detector. Spectrograph utilizes a linear array detector like charge coupled device (CCD) or photodiode array (PDA) in place of a single element detector and light signals at multiple wavelengths can be detected simultaneously. Diffraction gratings based instruments are relatively low in cost and very capable in many industrial sectors. Drawbacks for the scanning diffraction grating monochromators are relatively slow scanning speed, and degrading system performance over time due to mechanical fatigue

of moving parts. Compared to scanning monochromator, spectrograph is faster in speed, has no moving parts, and thus is robust in structure.

The other type of dispersive system employs electronically tunable filters, such as acousto-optical tunable filters (AOTF). Using AOTF as the dispersive device, spectrometers can be constructed with no moving parts, having very high scanning speed, wide spectral working range, and random wavelength access (Eilert 1991). Compared to diffraction grating spectrometers, the electronically tunable filter based instruments have much higher cost and thus are not widely applied yet.

Non-dispersive systems: There are mainly three groups of non-dispersive systems. The first group of spectrometers is based on the use of Fourier transform (FT) and Michelson interferometer. This type of instrument is mainly installed in research laboratories. The working principles have enabled the system to give excellent wavelength precision and accuracy, very high signal-to-noise-ratio (SNR), and relatively fast scanning speed. Since it utilizes a Michelson interferometer to create the conditions for optical interference by splitting light into two beams and then recombining them after a path difference has been introduced using a moving mirror, the system is very delicate and performance is subject to external mechanical vibration and dust.

The second group of non-dispersive system is based on a limited number of interference filters. It is the simplest and cheapest NIR instruments. Optical filters are usually chosen according to the absorption wavelengths used for the most popular applications, e.g. protein, moisture, and oil content in agricultural samples. Therefore, interference filter based instruments are only designed for a limited range of routine analyses.

The third group is the light emitting diode (LED) based instruments. This type of instrument employs an array of LEDs that emit narrow bands of NIR light as the illumination sources. As the emitting wavelengths are predetermined the instrument is usually dedicated to a specific series of measurement. Both LED and filter based instruments satisfy the needs for low cost, specific applications, and portable instrumentation for field analyses.

Generally, selection of appropriate instrumentation configuration depends on the purpose of application. More extensive review on instrumentation for vibrational NIR spectroscopy can be found in other references (Coates 1998; Osborne et al. 1993).

2.4.2 The spectrograph

For field applications, instruments based on fixed dispersive optics, e.g. diffraction grating or AOTF, and sensor arrays have proven to be a robust solution when multi-wavelength spectral data are required. There have been some reports on spectrograph system applied to optical spectroscopy and its comparison with other instrument setups.

Mayes and Callis (1989) described a spectrograph system working in the 600 - 1100 nm wavelength range and used it for absorption and diffuse reflectance measurement. The paper addressed the problems associated with scanning spectrometers and stated that a spectrometer based on multi-channel detector, like PDA detector, provided more rugged, compact, and reliable solution. The authors constructed a system employing a 150 grooves/mm holographic grating (contrary to ruled grating due to a different ruling process) optimized for the 700 - 1500 nm region as the dispersive device and a silicon PDA detector. The authors then tested several optical parameters of the spectrophotometer. A comparison was made between a commercial scanning

monochromator and their PDA spectrometer, and it was shown that their PDA spectrometer had superior performance.

Complementary to NIRS that studies absorption of matter, Raman spectroscopy studies the inelastic scattering of light by sample (Shreve 1952). The instruments for both spectroscopic studies are similar with some minor modifications. Williamson et al. (1989) reported a NIR Raman system based on Diode laser, diffraction grating, and CCD array detector. The authors employed an ISA-640 300 line/mm ruled grating coupled with a 512×512 CCD detector binned along one dimension to detect Raman spectra. An NIR Raman spectrum of 0.1 M $(NH_4)_2SO_4$ in water was acquired using the system with a high SNR of 300 at 1.2 cm^{-1} resolution. They finally concluded that the diode laser/CCD spectrograph configuration was most useful in routine or repetitive analysis because its advantages in price and ruggedness over Fourier Transform-Raman system.

2.4.3 Working principles of major devices in a spectrograph

This section briefly describes the theory of three major devices in a spectrograph system. An optical fiber bundle works on the principle of total internal reflection and is used to collect reflected light from sample. The diffraction grating is the core part of the optical system. It disperses polychromatic light into different spatial angles according to its wavelengths. A CCD device with a linear array of sensing elements can be configured to detect the spatially dispersed light from diffraction grating.

2.4.3.1 Diffraction grating and optical spectroscopy

As one of the most powerful analytical tools available to the chemists, optical spectroscopy conducts the elemental and molecular analysis of complex samples based on emission, absorption, fluorescence, and Raman effects which are related to wavelength composition of light. The diffraction grating can be visualized as a piece of glass with many parallel reflecting (or transmitting) elements on the surface. These elements are separated by a constant distance and its inverse is called groove density or grooves per millimeter. Since the diffraction grating can disperse polychromatic light into monochromatic light according to its wavelengths, the device is of considerable importance in spectroscopy. Strong (1960) quoted "It is difficult to point to another single device that has brought more important experimental information to every field of science than the diffraction grating. The physicist, the astronomer, the chemist, the biologist, the metallurgist, all use it as a routine tool of unsurpassed accuracy and precision, as a detector of atomic species to determine the characteristics of heavenly bodies and the presence of atmospheres in the planets, to study the structures of molecules and atoms, and to obtain a thousand and one items of information without which modern science would be greatly handicapped".

2.4.3.2 The grating equation

As mentioned earlier, the diffraction grating can be viewed as a collection of straight, parallel, and equally spaced grooves on a plane (for plane gratings) or curved optical surface (for concave gratings). Each groove can be thought of as a thin slit that provides diffracted light. The diffraction grating operates by the mutual interference of light beams diffracted by the grooves. Figure 2.3(a) demonstrates the diffraction of monochromatic

light by a reflection plane grating. When a light ray of wavelength λ strikes the grating surface at an angle α, the light is diffracted by the grating along angles β_m. The incident angle α and diffraction angles β_m are measured from the grating normal. Signs for angles depend on the sides the diffracted light takes with respect to the incident light. In Fig. 2.3(a), $\alpha > 0$ and $\beta_1 > 0$ while the angles $\beta_0 < 0$ and $\beta_{-1} < 0$. The equation that governs operation of diffraction grating is given by:

$$m\lambda = d(\sin\alpha + \sin\beta_m) \tag{2.13}$$

where m is the diffraction order (no unit), incident angle α (degree), diffraction angles β_m (degree), wavelength λ (m), and grating constant d (m).

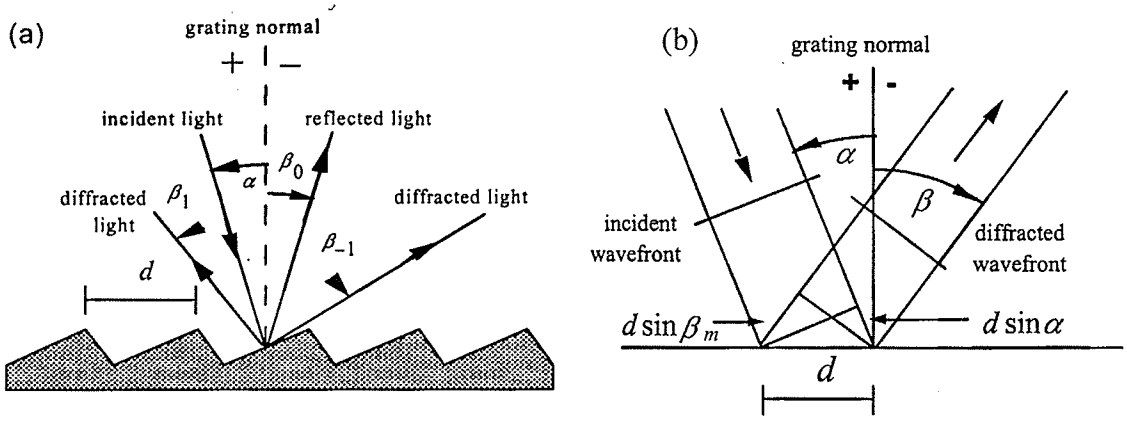


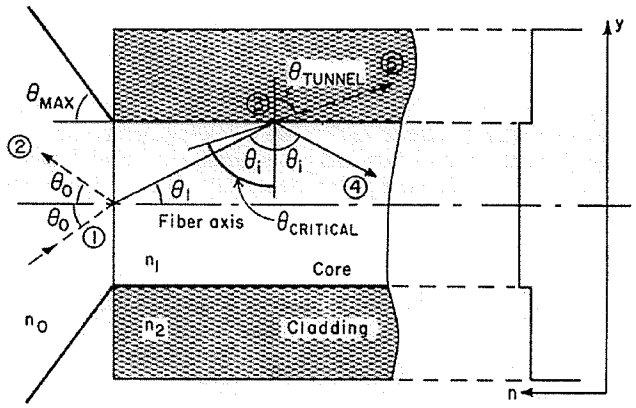
Figure 2.3 Diffraction of a beam of monochromatic light by a reflection plane grating (a) Incident light at angle α is diffracted into angles β_0, β_1, and β_{-1}. Angles are measured with respect to grating normal; (b) Geometry of diffraction for planar wavefronts. *Source:* Palmer (2002).

The equation can be easily visualized by the relation shown in Fig. 2.3(b). Two planar wave fronts are diffracted by the optical surface of plane grating, and introduce a path difference that equals $d(\sin\alpha + \sin\beta_m)$. When the path difference equals an integer

of wavelengths, two portions of light will interfere constructively. Sign convention for the angles is to be observed in calculation, *i.e.*, $\beta_m < 0$ in Fig. 2.3 (b) and $m < 0$.

2.4.3.3 Fiber optical waveguide

The fiber optical light guide is an important part in spectrometer design since it allows great flexibility in instrument design and sample presentation. The phenomenon of total internal reflection at the interface between two dielectric media can explain the working principle of an optical fiber. Typical optical fibers are solid cylinders. The step index fiber and its meridional ray paths in a step index fiber are shown in Fig. 2.4.



(a) (b)

Figure 2.4 Step index fiber and transmission curve (a) Meridional ray paths in a step index fiber; (b) Step refractive index profile. *Source:* Allard (1990)

As illustrated in Fig. 2.4(a), an optical fiber consists of a core and its cladding (the shadowed parts). The incoming light strikes the core and air interface at θ_0 to the fiber axis, and part of the ray is transmitted into the fiber. Then this portion of light ray strikes the interface of core and cladding and is refracted again. If θ_i is smaller than the $\theta_{CRITICAL}$, portion of the light propagates into the cladding and eventually escapes from the fiber. Numerical Aperture (NA) is defined as the sine value of the largest acceptance angle into

the fiber that a transmitted ray has an incident angle θ_i greater than $\theta_{CRITICAL}$. The critical angle $\theta_{CRITICAL}$ and NA are given by:

$$\sin\theta_{CRITICAL} = \frac{n_2}{n_1} \tag{2.14}$$

$$NA = \sin\theta_{\max} = (n_1^2 - n_2^2)^{\frac{1}{2}} \tag{2.15}$$

where NA is the numerical aperture (no unit), $\theta_{\max}$ is the maximal acceptance angle (degree), n_1 is the refractive index of optical fiber core (no unit), and n_2 is the refractive index of the cladding (no unit).

2.4.3.4 Physics of charge coupled device image sensor

In its simplest form, a linear CCD is an array of closely spaced metal-oxide-semiconductor (MOS) capacitors. The CCDs are usually fabricated in a p-type silicon substrate with a layer of silicon dioxide (SiO_2) grown onto it. The metal electrode, which has connection with the wires, is then evaporated on top of the oxide layer. When illuminated with radiation, charge packets can be generated and trapped in potential wells under those electrodes with positive voltages. The trapped charge is proportional to the total incoming photons that strike the silicon substrate during the integration period. The transfer of charge from the collection site to the output node of the CCD array is accomplished by shifting the charge from one element to its adjacent one. Figure 2.5 illustrates the three-phase CCD detector utilizing three clock waves to control transfer.

The basic principle of transferring trapped charge is to use three voltage lines $L_i (i = 1, 2, 3)$ with each line connected to every third electrode G_i ($i = 1, 2, 3$). Assuming that initial potential of L_1 is positive, while L_2 and L_3 are kept at zero potential, element

underneath G_1 will then trap charges generated by incoming light. After integrating the signal for a period, potential of L_2 is raised, and the trapped charges under electrodes G_1 are shared between G_1 and G_2. By reducing the voltage applied to line L_1, trapped charges are gradually transferred to element under G_2. Repeating the above process eventually shifts the trapped charge packets to the output electrode.

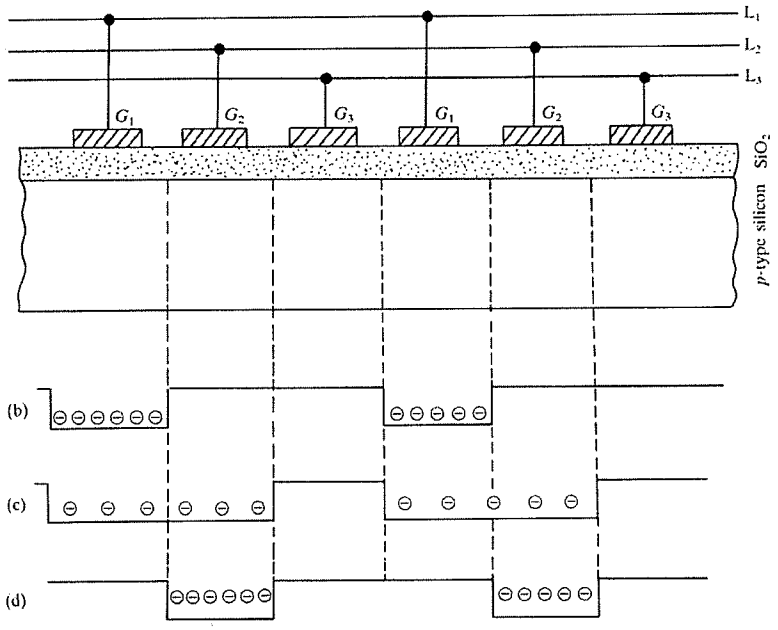


Figure 2.5 Structure of a three-phase CCD array detector and charge transfer process. *Source:* Wilson and Hawkes (1989).

2.5 Chemometrics

Workman et al. (1996) gave a concept of chemometrics stating that "Chemometrics can generally be described as the application of mathematical and statistical methods to 1) improve chemical measurement processes, and 2) extract more useful chemical information from chemical and physical measurement data." Handling spectral data can

be divided into two steps: spectral data preprocessing and exploratory or quantitative analysis of transformed data.

2.5.1 Spectral data preprocessing methods

Preprocessing methods and data transformations alter spectral data according to the characteristics of the entire data set. There is a wide array of data preprocessing algorithms available. These methods have been developed to smooth and correct baseline shift, to improve SNR of spectra, to compress spectral data, to correct for nonlinearity problems due to light scattering effect, etc.

2.5.1.1 Baseline correction, smoothing, and data compression techniques

Spectroscopists can often find that spectral response ride on a changing signal with constant slope that represents normal background level, which is referred to as baseline. Taking the first or second derivatives of the NIR spectra can remove spectra baseline. Methods to calculate derivatives like Savitzky-Golay method (Savitsky and Golay 1964) or Gap method (Norris and Williams 1984) can as well smooth the data using data from a spectral window. Savitsky and Golay proposed the filtering method to smooth spectral data and calculate derivatives, which became a standard preprocessing method in today's chemometrics. The first derivatives are often used to remove baseline effects and the second derivatives correct for baseline offsets and sloping baselines. Holler et al. (1989) applied second derivatives in curve fitting and found that the method was effective in eliminating redundant information. The drawback as commented by the authors was that there is a possibility of over-smoothing overlapping peaks. Besides Savitzky-Golay

method, FT techniques and other low-pass filtering techniques like binomial smoothing could also be used to smooth spectra.

Fourier transform and principal component analysis (PCA) techniques could both be used to compress the dimensionality of original data and smooth the spectra (Brereton 2003). Both methods retain a significant amount of information present in the original data using much fewer variables, which are referred to principal components for PCA and frequency components for FT. McClure and Davies (1988) were the first to examine mathematics behind fast Fourier transform (FFT) and proposed the instrumental application of FFT for spectra smoothing and data reduction. Chau et al. (2004) reviewed basic and concurrent techniques in the field of chemometrics and described how wavelet transform (WT) can be used in smoothing, denoising, compressing and enhancing the resolution of spectral data. The potential of WT in chemometrics has been corroborated by a number of publications on application of wavelets in spectroscopy (Ren and Gao 2000; Shao et al. 2000).

2.5.1.2 Light scattering correction techniques

Since most cereal grain products are solid and opaque. Most of light projected onto these solid samples are scattered. Spectral measurement of most cereal food is therefore conducted under diffuse reflectance mode. However, it was pointed out that a significant portion of the reflectance signal were direct results of light scattering noise (Martens et al. 2003). To reduce calibration complexity, this scattering noise can be corrected using methods like standard normal variate (SNV) followed by detrending (DT) (Barnes et al. 1989) or multiplicative signal correction (MSC) method (Geladi et al. 1985).

Barnes et al. (1989) cited the SNV transformation from statistics to eliminate variation due to particle size. Each spectrum was corrected according to the following equation:

$$SNV_i = \frac{y_i - \bar{y}}{\sqrt{\frac{\sum(y_i - \bar{y})^2}{n-1}}} \tag{2.16}$$

where SNV_i is the individual standard normal variation for ith wavelength (no unit), y_i is the spectral data at ith wavelength variable (no unit), $\bar{y}$ is the mean of ith wavelength variable (no unit), and n is the total number of samples.

In their experiment, the authors used sucrose in crystalline and finely powdered form and a set of grass hays to illustrate the transformation. Principal component analysis was carried out on spectra before and after the SNV transformation and it was found that SNV effectively removed the portion of variance explained by the first principal component, which had little structure before applying SNV transformation. The authors also indicated that SNV could increase the correlation between NIR spectra with the component concentration determined in laboratory. The authors commented that there existed a linear trend of increase in reflectance value levels over the spectral range of 1100 to 2500 nm. The DT method is a second-degree polynomial curve-fitting technique employed to correct for the curvilinear trend of spectral response over wavelength. For densely packed samples, the linear trend became curvilinear and a second-degree polynomial proved to be effective in correcting this trend. Calibration results revealed that SNV and DT achieved calibration statistics comparable or better than that obtained by taking first or second derivatives of log (1/R) spectra.

Geladi et al. (1985) proposed the MSC approach to separate light scatter from absorption in NIR diffuse reflectance spectra. Multiplicative signal correction attempted to estimate two coefficients, a and b, that ideally contain all the physical information in diffuse reflectance spectra based on a linear regression model (Pedersen et al. 2002).

$$X = a + b\overline{X} + E \tag{2.17}$$

where X is the diffuse reflectance spectrum of the sample (no unit), and $\overline{X}$ is the spectrum of the "ideal" sample, which is often the average spectrum(no unit). The residual spectrum about the MSC model, E, was assumed to ideally represent the chemical information in X (no unit). Once a and b were estimated, the corrected spectrum was calculated according to:

$$X_{corrected} = (X - \hat{a})/\hat{b} = \overline{X} + \hat{E}/\hat{b} \tag{2.18}$$

where $\hat{a}$, $\hat{b}$, and $\hat{E}$ are estimate of a, b, and E, respectively. None of these variables have any units.

Their experiments utilized grounded bovine and porcine samples with fat content to be determined. Three sample sets were studied to examine the effect of physical difference (homogeneity and surface smoothness) on spectral variance, linearity of instrumental response, and optimal model complexity for different data transformations. It was pointed out that the working basis of MSC was the difference between wavelength dependence of light scattering and absorption. The MSC algorithm was examined for its ability to remove scatter effect and linearize spectral data. Their study of the regression coefficients in MSC in relation to fat content indicated that there was strong correlation for multiplicative term with that of fat content. The paper summarized that MSC could

provide more parsimonious calibration even though partial least-square regression (PLSR) could handle the nonlinearity problem by including more factors in the calibration.

A later paper by Isaksson and Næs (1988) examined the effect of MSC to improve the prediction ability of linear regression models. A comparison to other research works was furnished and a favorable conclusion towards MSC was drawn. Piece-wise multiplicative scatter correction (PMSC) was later proposed by Isaksson and Kowalski (1993) and their research findings indicated a favorable opinion of PMSC over MSC in terms of prediction performance. Azzouz et al. (2003) made a comparison between SNV and DT, first and second derivatives, MSC, and orthogonal signal correction (OSC) for NIR reflectance spectra, and it was found that SNV combined with DT produced good prediction accuracy and figure of merits on average. Generally, the performance of each preprocessing method depended to a great extent on the application and the characteristics of the spectral data.

2.5.2 Quantitative analysis techniques

There have been many choices of regression techniques in NIR spectroscopy. Mark (1991) described the application of multiple linear regression (MLR) to NIR calibration in great depth. Cowe and McNicol (1985) employed principal component regression (PCR) technique in their analysis of NIR spectra. The most successful quantitative analysis technique was partial least square regression (PLSR) proposed by Martens and Næs (1989).

Multiple linear regression (MLR): Multiple linear regression is one of the most widely used statistical techniques. It utilizes variables referred to as predictor variables (e.g., absorbance at different wavelengths) to infer the response variables (e.g., different

chemical compositions). Assuming there are m predictor variables (x) that could be measured easily compared to p response variables (y). Measure n samples and one can get a predictor matrix **X** of size n by m, and a response matrix **Y** of size n by p. Multiple linear regression tries to find a first order linear relation between these two matrix according to:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E} \tag{2.19}$$

where **B** is the regression coefficient matrix (no unit), and **E** is the residual matrix (no unit).

Principal component regression (PCR): There are two steps in PCR technique. The first step involves matrix decomposition using PCA, and the second step involves regression using the score matrix **T** in place of the original data matrix. Score matrix has better properties than the original data matrix because the score vectors are orthogonal to each other and span the original space of **X**. Score matrix derived using PCA and regression coefficients can be calculated according to:

$$\mathbf{T} = \mathbf{XP} \tag{2.20}$$

$$\mathbf{Y} = \mathbf{TB} + \mathbf{E} \tag{2.21}$$

where **P** is the loading matrix in PCA (no unit), and **B** is the regression coefficient matrix (no unit), and **E** is the residual matrix (no unit).

Partial least square regression (PLSR): The PLSR algorithm is described as a full spectrum calibration method. It combines the two stages of data decomposition and regression together, with the resulting PLS factors describing variation both dominant in original data matrix as well as relevant for regression task. The PLSR algorithm improves the inner relation of spectral data matrix and chemical data matrix in calibration by

exchanging scores during their decomposition using nonlinear iterative partial least square (NIPALS) method.

In a tutorial on PLSR, Geladi and Kowalski (1986) made a comparison among available regression techniques, MLR, PCR, and PLSR, and illustrated the fundamental ideas behind PLSR technique. A discussion of MLR regarding situations when $m < n$, $m = n$, and $m > n$ was put forward. It was shown that variables have to be deleted when $m < n$, and there may be a problem in MLR called collinearity when $m = n$, and $m > n$. Principal component regression (PCR) fixes the MLR problems by representing the original data matrix by its score matrix. When performing PCR, data decomposition by PCA and calibration regression is carried out independently. Therefore, PCR only considers the most dominant dimensions in original spectral data matrix but not the ones most relevant for determination of chemical data. The PLSR algorithm improves the PCR by seeking partial least square (PLS) factors that are both dominate in the independent variable matrix **X** and dependent variable matrix **Y**. By including appropriate number of components, the method can model some of the nonlinear effect. A discussion on choosing the number of components in the PLS model was put forward. The NIPALS and PLSR algorithms were included in the paper.

In recent years, ANNs, due to their ability to model almost any nonlinear relationship, have attracted the interest of spectroscopists. Research activities in using ANN for NIR spectra analysis have resulted in a number of publications (Ding et al. 1999; Thodberg 1996).

2.6 Near-infrared Spectroscopy in Agriculture

In agricultural industry, NIRS was used to quantitatively determine the moisture content in soybean flour as early as 1960s (Ben-Gera and Norris 1968) but it took a few more decades before NIRS was widely accepted. Today, with advancement in NIRS instrumentation and maturity of mathematical calibration techniques of NIR spectra, NIRS has become the method of choice for routine determination of physical and chemical properties of some of the cereal grain products. A brief review of the recent application of NIRS in agriculture and its application in determination of chemical content of ground wheat is provided.

2.6.1 General application of NIRS in agriculture

In the period that followed year 2000, there have been quite a number of papers on using NIRS in the agricultural and food industry. The reported research outcomes covered a wide spectrum of topics revealing the great potential of NIRS.

Archibald and Kays (2000) examined the potential use of NIRS for intact food quality analysis. The authors collected spectra of 136 assorted cereal food products with a DA7000 spectrometer. The authors preprocessed spectra using second derivatives followed by SNV and mean-centering. The spectral window with the lowest root mean square error of cross-validation (RMSECV) was chosen for calibration of intact samples without milling. Final results revealed that although there was a degradation of prediction performance for intact samples, it was demonstrated that the NIR reflectance spectroscopy performance was overall useful considering the 49.4% range of total dietary fiber (TDF) in diverse cereal food products. Near-infrared reflectance spectroscopy, thus, was suggested as a fast and non-destructive method for intact food product analysis.

Wesley et al. (2001) measured gliadin and glutenin contents, which are the major compositions of protein in flour, and then used the information to determine the quality of protein. The paper tried to build calibration models for protein composition that were independent of total protein contents. Curve-fitting methods and PLS were compared on a set of samples with wide range of physical properties and cultivars grown at different locations across Australia. It was shown that PLS performed favorably compared with curve-fitting but still there existed an unacceptably large error. However, the authors suggested that such coarse estimate was acceptable in plant breeding situations where classification would be done on the basis of high or low glutenin content. Despite the large bias between actual and predicted values by curve-fitting, the paper indicated that the curve-fitting method could sufficiently rank the components into high, medium, and low for each component independent of total protein content.

Fontaine et al. (2002) achieved accurate and fast prediction results for methionine, cystine, lysine, threonine, tryptophan, and other essential amino acids, protein, and moisture in wheat, barley, corn, triticale, wheat bran/middlings, rice bran, and sorghum. It took the authors over five years to collect 1100 samples of global origin. Amino acids and other chemical contents were determined using chemical and chromatographic methods. The NIR spectra were then collected using a composite monochromator and calibrations were developed and validated using independent test sets. Satisfactory calibration statistics for each kind of cereal products were provided. The calibration models could be transferred with good accuracy to other NIR spectrometers and international labs. The authors suggested that successful application of NIRS can help to

improve the accuracy of feed formulation and obtain better quality and lower production costs.

Xie et al. (2003) found NIR spectra correlated with the actual bread storage time better than Texture Analyzer (TA), which only measures bread firmness. The reason was that NIRS captures both the physical and chemical changes. Meanwhile, it was shown that NIRS was less sensitive to batch difference. Repacking samples (5-6 times) and keeping the same sample orientation played important role in their spectral measurement.

Dowell and Maghirang (2002) reviewed current single kernel characterization systems and sorting systems that addressed the limitations of bulk-sample analysis. Besides physical and chemical information, the single kernel characterization system SKCS4170 (Perten Instruments AB, Sweden) coupled with a Si-InGaAs diode-array spectrometer (400-1700 nm) could render information regarding content distributions, fungal damage, and insect infestations without destroying samples but at relatively low speed (two kernels per second). High speed sorting systems were not able to quantify wheat attributes. The authors suggested that single kernel characteristics as measured by NIR spectroscopy had twice as large a standard error as that of bulk-sample analysis. The system was unable to detect trace elements present in cereal grains. Meanwhile, cost and speed have restrained its acceptance in grain industry.

Maghirang et al. (2002) conducted a two-month research using NIRS to detect insects in wheat insects. Their findings revealed that factors influencing NIRS calibrations for live and dead insects were generally the same. Important wavelengths for detection and correct classification rates were studied. The authors suggested that

calibrations could be done after killing wheat insects and be transferred without adjustment. A low detection rate for small larvae was achieved in this study.

Dowell (2000) presented NIRS as an objective measure for classifying vitreous and non-vitreous durum wheat. Test results agreed with subjective classification perfectly when samples were obviously vitreous and non-vitreous. When dealing with difficult-to-classify samples, NIRS provided quantitative measures to reduce inter-inspector variability. Basic understanding of classification was provided on basis of the authors' experience and is not theoretically supported.

Wang et al. (2001) assessed heat damaged wheat kernels using NIRS. The paper discussed current inspection methods and several other methods to assess heat damage in wheat. The authors pointed out that visual examination of wheat kernels based on color changes was subjective and depended on experience. Other methods required expensive equipments and were time-consuming. The paper studied NIR reflectance spectroscopy and stated that NIRS met the requirements of a rapid, precise, and non-destructive test system based on single kernels. They employed a diode-array NIR spectrometer in their study. The experimental results indicated that PLS models using spectra in the region of 750 - 1700 nm rendered 100% classification accuracy of heat-damaged and undamaged wheat kernels. The authors assessed a two wavelength model that gave a classification accuracy of 96.8% for two NIR wavelengths and 94.3% for two visible wavelengths. The authors recommended that the simple two wavelength model be applied in industry and grain inspection applications.

Paliwal et al. (2004) demonstrated how NIRS could be used to differentiate between *Sitophilus oryzae* (rice weevil) and *Rhyzopertha dominica* (lesser grain borer) at

their pupal stage at different infestation levels. The authors employed PCA to analyze spectra of wheat kernels infested with two types of insects. Their finding stated that the difference between two groups of spectra in the score space became obvious as the infestation levels increase. The authors also tried to quantitatively determine the infestation levels of rice weevil at egg, larval, pupal, and adult stages. The data were preprocessed using MSC and SNV followed by DT. Regression was done using PLSR. The outcomes indicated that both MSC and SNV followed by DT are effective in removing irrelevant scattering signal. Quantifying infestation levels is less accurate at lower infestation levels compared with higher ones.

2.6.2 Properties of ground wheat as determined by NIRS

Hruschka and Norris (1982) tried to determine the protein and moisture content of ground wheat using least-squares curve fitting of NIR spectra. The authors assumed that the spectrum of ground wheat could be approximated by a linear combination of its component spectra, either using spectra of chemical constituent sets or sample components. The chemical component sets included reflectance spectra of starches, proteins, cellulose, and sugars, and the absorbance spectra of water, oils, and atmospherics. The chemical component method failed to produce a flat residual spectrum and thus was not further attempted. The sample components method selected seven to ten wheat spectra as components and provided flat residual spectra. It was found that the curve fitting method possessed better performance than the derivative technique for ground wheat analysis using NIRS. Moisture measurement results using quotient method were not reported in details.

Law and Blakeney (1996) examined the transformation of reflectance spectra using Kubelka-Munk function F(R) and apparent absorbance log (1/R). They applied both transformation methods to reflectance spectra of ground wheat to determine the moisture content. The authors discovered that moisture content in ground wheat could be determined directly using Kubelka-Munk function applied to reflectance at 1935 nm and 2100 nm and achieved a standard error of prediction (SEP) of 0.30. They tested ground wheat of a wide range of wheat types, particle size and moisture content and believed the calibration could be successfully transferred among different instrument types.

Pasikatan et al. (2001) stated the importance of automatic inspection of flour milling and studied the use of NIRS to determine the particle size of ground wheat. The authors examined the effect of wheat class and roll gap on particle size distribution. They collected reflectance spectra of ground wheat of four fractions of particle sizes using a diode array spectrometer. The experimental results indicated that preprocessing of spectra using unit-area combined with baseline correction and second derivative yield best models. For each of the four size fractions (> 1041, > 375, > 240, > 136 μm), the best models had SEP values of 4.07, 1.75, 1.03, and 1.40 and r^2 of 0.93, 0.90, 0.88, and 0.38, respectively. They also stated that granulation sensing using an NIR reflectance spectrometer became more accurate as particle size became finer for succeeding break and reduction stages.

2.7 Some Moisture Measurement Methods

Despite the efforts that went into investigation of NIR spectroscopy as a viable technique for moisture content measurement in wheat (Law and Tkachuk 1977a; Williams 1975; Williams and Thompson 1978) and other solids (Law and Tkachuk 1977b; Stermer et al.

1977), several other methods have been used to determine the moisture content in cereal grains either as basic reference methods or other practical methods. The approved methods by American Association of Cereal Chemists (AACC) are electrical methods and air oven methods (standard reference). Besides NIRS, nuclear magnetic resonance (NMR) spectroscopy is another possible but quite expensive method for moisture measurement.

Electronic moisture meters are based on the electrical conductance and capacitance properties of grains, which are directly related to the moisture content of grains. In the United States, the capacitance-type meters have been exclusively used in grain marketing. However, the method still needs much improvement in its measurement accuracy and robustness. These instruments are calibrated using samples of clean, uniform, and high-quality grain, but in a practical scenario grains usually contain varying amounts of weed seeds, broken kernels, and foreign materials. Therefore, factors like sample adulteration, growing season, high moisture content, and instrument difference may all contribute to inaccurate measurements when using these electronic devices.

Air oven methods have two standard procedures, that is, one-stage and two-stage procedure. The one-stage procedure involves using an oven capable of maintaining a temperature of 130°C (±1°C) uniformly throughout oven and providing good ventilation. The measurement includes grinding 30 to 40 g samples using a laboratory mill, weighing 2 to 3 g samples to be dried, and calculating the weight loss after 60 min heating in the oven with a constant temperature at 130°C. The two-stage procedure is intended for moisture content measurement higher than 16%. This is because excessive amount of moisture content will be lost when grinding samples of high moisture level. The

procedure require that whole grain sample to be weighted and dried in a well ventilated place till the moisture reaches air-dry condition at the first stage. The moisture lose is then calculated and the air-dried sample is tested following the same procedures as those in the one-stage method described above. Details could be found in AACC method 44-15A (AACC, 1983)

Determination of moisture content in cereal grain using NMR is not widely accepted due to the prohibitive cost of equipment. The NMR signal must be calibrated against moisture content as given by reference method such as air oven method (Christensen et al. 1992).

3. MATERIALS AND METHODS

3.1 System Design

The entire spectrographic system consisted of an optical dispersive system, a photon detection system, a data acquisition system, and a virtual instrument (VI) interface program. The optical dispersive system was composed of a reflection diffraction grating, lens, baffles, an optical fiber probe, and mechanical mounts for these components. A linear CCD array detector with an output range of 0 - 5 V was placed at the focal plane of the imaging lens. The data acquisition system was set up to acquire CCD detector output with different number of samples corresponding to different number of wavelength variables. The data acquisition was synchronized using external triggers generated by the CCD device. Programs for monitoring, spectral data acquisition, and data logging were developed under LabVIEW environment.

3.1.1 Optical system

The optical system was mounted on a 300 mm × 300 mm optical bench plate and sealed in a lightproof tank. A fiber optical light guide was mounted through a pinhole in front of the spectrograph system to collect reflectance signal from sample to be analyzed. As the core device in the optical dispersive system, the plane reflection diffraction grating needed ancillary optics to collimate light from fiber optical waveguide and focus diffracted light that was dispersed into different angles according to the grating equation.

3.1.1.1 Selection of diffraction grating systems

When selecting a diffraction grating for specific applications, consideration should be given to factors such as cost, performance, design requirements, and availability of the product. Some rules in choosing the appropriate diffraction grating for the optical dispersive systems were:

Reflection versus transmission: Reflection grating systems were much more common than transmission grating systems. The reflection grating systems could provide more compact system size and were not restricted by the transmission properties of the substrate.

Plane reflection versus concave reflection: Plane reflection gratings were preferred over concave gratings in the infrared region because of the availability of large gratings. The much more expensive concave reflection gratings were suitable for ultraviolet region where focusing optics was hard to find.

The specifications for the selected diffraction grating are listed in Table 3.1.

Table 3.1 Specification table for the purchased diffraction grating

Coating	Gold
Dimensional Tolerance (mm)	±0.5
Thickness (mm)	9.5
Material	Float Glass
Groove Direction	Parallel to Short Dimension
Size (mm)	30 x 30
Design Wavelength (nm)	1000
Grooves per mm	1200
Blaze Angle	36° 52'
Stock Number	NT55-263

In this study, a plane reflection grating system with performance optimized for NIR region was selected. The ruled gold-coated versions of our commercial gratings provide a 15-20% increase in absolute efficiency in the 0.7 to 1.1 microns spectral region. This made them an excellent choice for our application that involved spectroscopy setups using silicon array detectors. The physical dimension of the grating was 30 mm × 30 mm with a groove density of 1200 grooves per millimeter. The designed wavelength is 1000 nm with a blazed angle of 36° 52'. In the spectrograph system to be constructed in this study, the diffraction grating operated in the first order which meant the diffraction grating had maximal free spectral range in its first-order application:

$$F_\lambda = \Delta\lambda = \frac{\lambda_1}{m} \tag{3.1}$$

where F_λ is free spectral range (m), λ_1 is wavelength of light diffracted in order m (m), $\Delta\lambda$ is wavelength difference between light diffracted along the direction λ_1 and of adjacent order $m+1$ (m).

This eliminated the need for filters to remove interference from light of overlapping spectral orders. For detailed discussion of diffraction gratings, more information could be found in the handbook by Palmer (2002).

3.1.1.2 Fiber optical waveguide

The fiber optical light guide employed in this project was a 300 mm long fiber bundle with a NA of 0.66 and acceptance angle of 82° (n_1 = 1.602, n_2 = 1.487). High transmission glass fibers were sheathed in PVC covered monocoil. The light guide ends were ground and polished with stainless steel end fittings. Due to intrinsic fiber losses, transmission efficiency could not reach 100% and the efficiency values varied around 60% with respect to wavelengths (Fig. 3.1). The transmission curve within spectral range

of 800 to 1100 nm was smooth. Therefore, it was apparent that there was no sharp absorption due to optical fiber losses within the short NIR range.

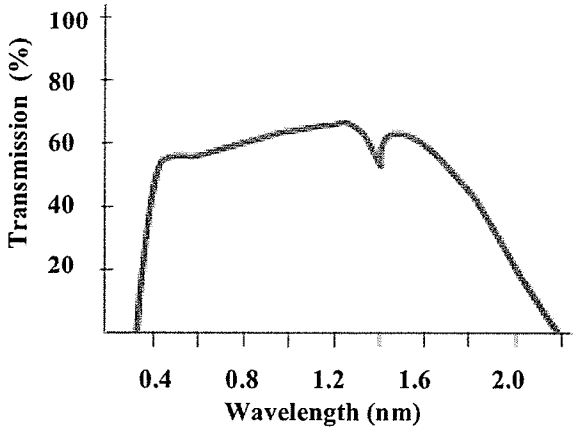


Figure 3.1 Transmission curve of the fiber optical waveguide. *Source*: Edmund Industrial Optics (2004)

3.1.1.3 The spectrograph optics

To efficiently utilize the optical signal collected by fiber optic waveguide, reduce effect of instrumental stray light, and properly focus diffracted light, light collimating lens system and imaging optics for the linear array detector were set up according to calculation in this section. A schematic of the optical system is shown in Fig. 3.2.

The functions, mechanical dimensions, and optical parameters for each component were as follows:

The aperture was used to constrain propagation of off-axis light that gave rise to stray light in the housing unit. The diameter (d) of aperture A was 2 mm. The housing unit for the entrance optics was used to shield stray light from propagating into the detection system.

The condenser lens L_1 had a diameter of 9.5 mm, was aberration corrected, and possessed a high light condensing efficiency. The lens (L_1) had an effective focal length (EFL) of 10 mm.

Lens L_2 was a specially designed achromatic lens that could correct for chromatic aberration in the 700 - 1100 nm region. The lens was used to collimate condensed light beam into a parallel light beam to be projected onto the diffraction grating. Lens L_2 had a diameter of 25 mm (D_2) and EFL of 40 mm.

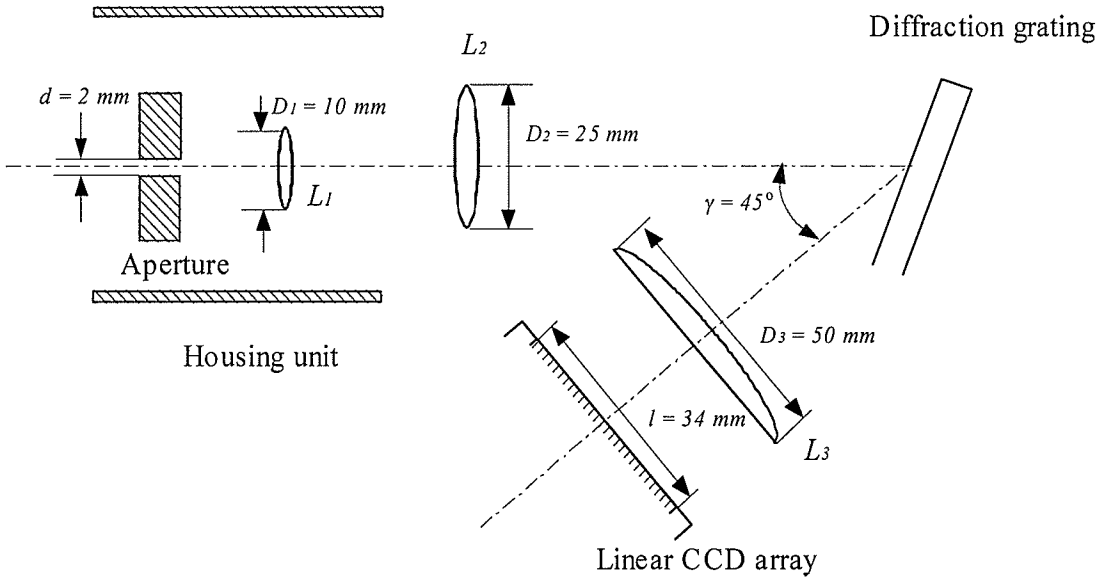


Figure 3.2 Optical system of the spectrograph.

Lens L_3 was a visible-NIR designed plano-convex (PCX) lens used to image diffracted light from the reflectance grating onto its focal plane where the linear CCD array detector was placed. The PCX imaging optics helped to transform the angular dispersion of light of different wavelengths into spatial dispersion along the CCD array detector. The coating of lens was optimized for minimum reflection in the NIR region. Lens L_3 had a diameter of 50 mm (D_3) and EFL of 50 mm. The width of the active area of

the linear CCD array detector was 36 mm (w), which corresponded to 2048 sensing elements.

Selection of these optical parts was based on theoretical calculations and availability of commercial optical components. Consideration was given to a simple optical system with low level of stray light and high system throughput. The diagrams of optical system are given in Fig. 3.3. Simple lens system was assumed in the theoretical calculation of the optical geometry.

The aperture A helped to reduce the etendue, which was the product of NA and the effective area of fiber optic waveguide at the output end, of input fiber optic light guide, and made it easier to match the incoming light to the entrance optics. The condenser L_1 with its optical mount served as the entrance optics to the spectrograph. L_1 helped to gather incoming photons, provided even illumination, and matched the f/value of the spectrograph. The f/value is given as:

$$f/value = \frac{1}{2NA} \tag{3.2}$$

where NA is the numerical aperture (no unit), which is the sine value of half angle.

In this project, a slitless spectrograph was constructed. This was mainly due to the fact that NIR spectra within this region are featureless, and absorption bands are fairly broad. It was illustrated in the results and discussion section when comparing spectra acquired using the spectrograph with spectra acquired using a spectrometer of higher resolution. Moreover, fabricating an entrance slit of very small size (generally on the order of microns to achieve high spectral resolution) was difficult, and defects in slit shape, dirt, and grease could adversely affect the system performance by degradation of spectral resolution.

When illuminating the spectrograph, it was often desired to under-fill the spectrometer optics to prevent stray reflections inside the spectrometer. This is illustrated in Fig. 3.3(a). The f/value$_{out}$ of entrance optics to the left side of spectrometer entrance needed to match the f/value$_{in}$ of the spectrometer, which referred to the accepting cone in front of the collimating lens. In the constructed spectrograph, optical aperture posed by L_1 was 6 mm (d_1) and it was decided to image the input aperture A at 15 mm (q), so as to get f/value$_{out}$ as 2.5 (q/d). The image was at the focal point of L_2 and the actual diameter of collimating lens filled by light was 16 mm ($d_2' = 40/2.5$). The theoretical resolving power of the system could be calculated as:

$$R = knW_g \tag{3.3}$$

where R is the resolving power (no unit), k the diffraction order (no unit), n the groove density (grooves/m), and W_g is the illuminated width of the grating (m).

When the system worked at an incident angle of 17°, the theoretical resolution was 20077 ($R = 1200\times16/\cos17^{\circ}$). The numerical bandpass (BP) at wavelength 1000 nm was calculated according to:

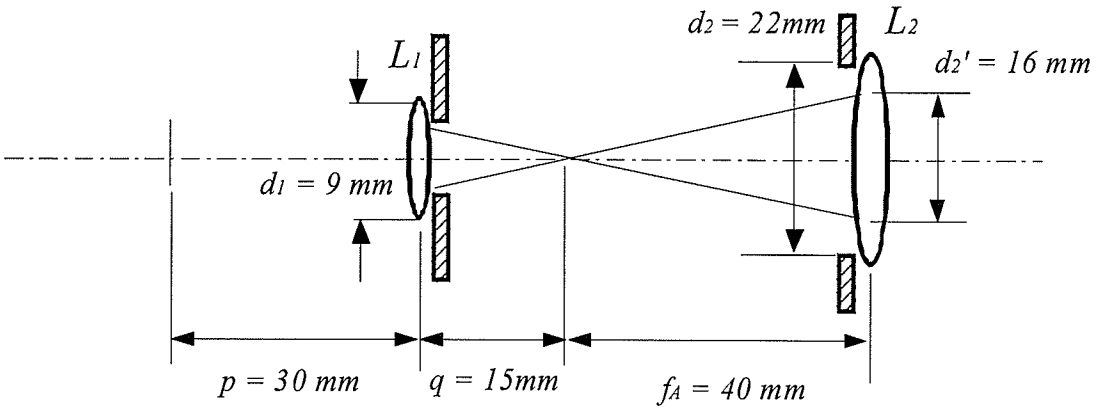
$$d\lambda = \frac{\lambda}{R} \tag{3.4}$$

The calculation gave a $d\lambda$ of 0.0498 nm (1000/20077). This value was the theoretically achievable BP if the entrance and exit slits were infinitely narrow and aberrations negligible. The actual instrumental profile characterized by the Full Width at Half Maximum (FWHM) was much broader.

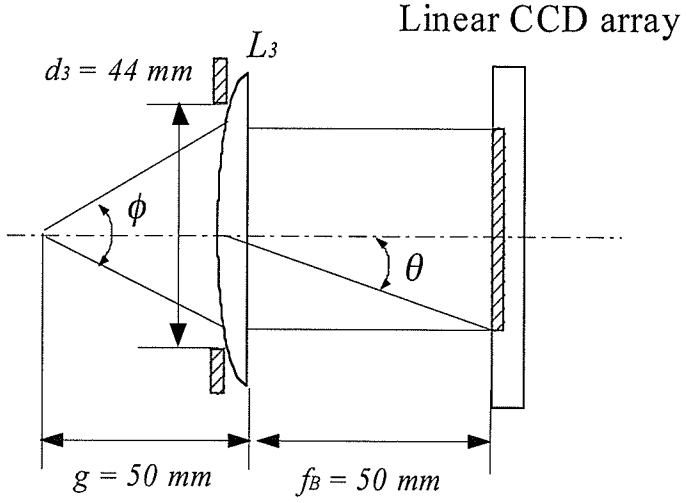
As shown in Figure 3.3(b), the minimal focal length available for a commercial PCX lens with a 50 mm diameter was 50 mm. Supposing that the light traveled along the

optical axis of L_3 to focus on the central pixel of the CCD detector, the maximal acceptance angle of the detection system could be calculated according to:

$$\theta = \arctan\frac{l}{2f_B} \tag{3.5}$$



(a)



(b)

Figure. 3.3 Theoretical optical system (a) entrance optics and collimating lens; (b) imaging optics for the CCD detector.

In this system, we had a θ of 18.8° (arctan(0.34)), and an acceptance angle of 37.6°. The real system might not be so perfectly aligned, and acceptance angle must be determined experimentally. It was important to note that there was vignetting of the spectral images at both ends of the CCD array detector output. This was because the spectrograph system was designed according to the light path along the input optical axis, and off-axis light was partially blocked by the plano-convex lens mount at those maximal acceptance angles and the lens aberration tended to smear the images.

In the spectrograph, the CCD array detector was placed on the focal plane of L_3. The dispersion and magnification factors varied across the focal plane. Assigning CCD array pixels to corresponding wavelength variables could be calculated according to the diagram shown in Fig. 3.4.

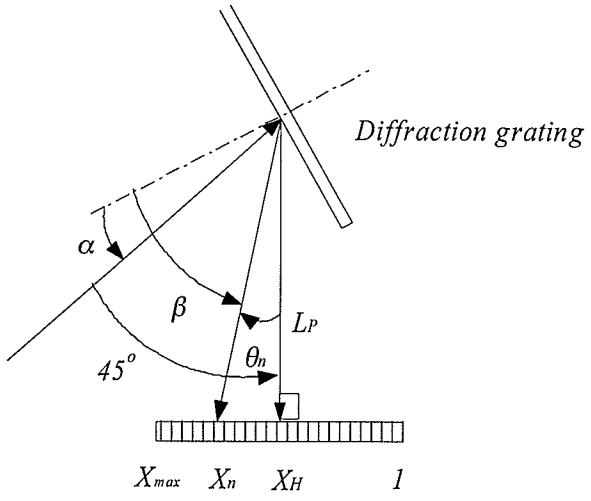


Figure 3.4 Relationship between wavelength and pixel position on the CCD array detector.

According to Eq. 2.13 and Fig. 3.4, diffracted light at wavelength λ and corresponding sensing element position X_n on CCD detector had the relationship:

$$n\lambda = \sin(45^\circ + \alpha - \arctan(\frac{X_H - X_n}{L_p})) + \sin\alpha \qquad (3.6)$$

where n is the groove density of the diffraction grating (grooves/m), L_p is the virtual focal length perpendicular to array detector (no unit), α is the incident angle to the diffraction grating (degree), β is the diffraction angle of light falling on position X_n (degree), θ is the angle between entrance optical axis and optical axis of L_3 (degree), X_H is the central position on array detector (no unit), X_{max} is the furthest position from X_H (no unit).

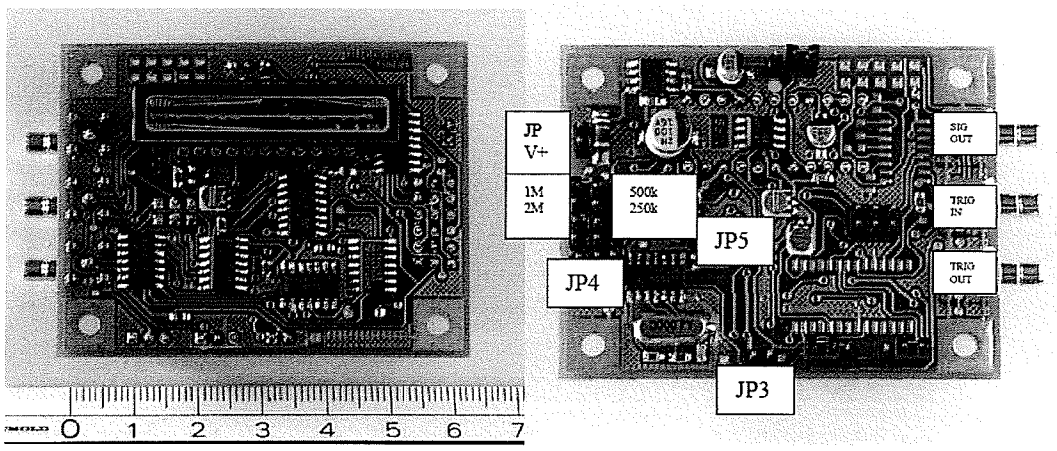
Details on theoretical design and configuration of diffraction grating spectrograph can be found in the book by Davis (1970).

3.1.2 Setting up the charge coupled device detector

A linear CCD-2000M array detector (Alphalas GmbH, Germany) was used in this study. The configuration settings of the device were presented in this section.

The CCD-2000M array detector employed in this study had 2048 pixels and worked in the spectral range from 320 nm to 1100 nm. The CCD array detector required a single 7.5 V DC power supply to operate. The manufacturer provided a set of jumpers to allow the detector to be configured to work in different modes. In this spectrographic system, the CCD-2000M array detector operated under auto trigger mode. The detector detected the incoming diffracted light from the diffraction grating within the integration period. Then these photo-generated charges were transferred to the output connector of the device. To synchronize data acquisition, the device continuously generated a series of voltage pulses to control the data acquisition of CCD detector output from the 2048 sensing elements.

The front and back view of the detector are shown in Fig. 3.5. A brief description of jumpers and connectors on the rear of CCD array detector was given in the following section. The configuration of each jumper in the spectrograph was also explained.



(a) (b)

Figure 3.5 The CCD2000M array detector (a) front side; (b) rear side. ***Source:*** **Alphalas GmbH (2004).**

Jumper **JP3** controlled the sampling mode between sample and hold (SH) and switched (SW) modes. When disconnected, the detector operated in SH mode. Signal coming from the detector pixels were integrated into a continuous output. In our application, we were concerned about the variation trend across pixels. Therefore JP3 was disconnected so that SH mode was selected (Appendix D).

Jumper **JP4** had multiple slots that controlled the scanning frequency. Four scanning frequencies, 0.25, 0.5, 1.0, and 2.0 MHz, were available. The lower was the scanning frequency, the longer was the integration time, and the higher sensitivity it could have under low light level. In the spectrographic system, the 0.25 MHz scanning frequency was selected to provide good sensitivity.

Jumper **JP5** was in charge of trigger modes. When disconnected, the device allowd external trigger. When connected, the device operated in auto-trigger mode. Jumper **JP5** was connected in this system to provide trigger signal to the data acquisition system.

Connector **JP V+** was the 7.5 V DC jack and was connected to a low noise stabilized 7.5 V power supply.

Connector **SIG OUT** was a sub-miniature, gold plated, clamp-type Sub-Miniature B (SMB) connector that output the voltage signal from CCD sensing elements (Appendix D).

Connector **TRIG IN** was a SMB connector that received external trigger signal when JP5 was disconnected. This connector was left open.

Connector **TRIG OUT** was a SMB that provided positive 5 V transistor-transistor logic (TTL) signal for approximately a one to 16 μs duration depending on scanning frequency (Appendix D).

Detailed timing features of the CCD array detector are listed in Table 3.2. The relation of these time features is shown in Fig. 3.6.

Table 3.2 Timing features of CCD array detector

Scanning Frequency (MHz)	T_0 (μs)	T_s (μs)	T_p (μs)	T_{delay} (μs)
0.25	14,280	8192	16	130
0.5	7143	4096	8	70
1	3570	2048	4	35
2	1800	1024	2	18

T_0 - Time between each internal trigger is generated
T_s - Total scanning time of all sensing elements
T_p - Pulse width of each generated trigger
T_{delay} - Time delay from trigger to first pixel

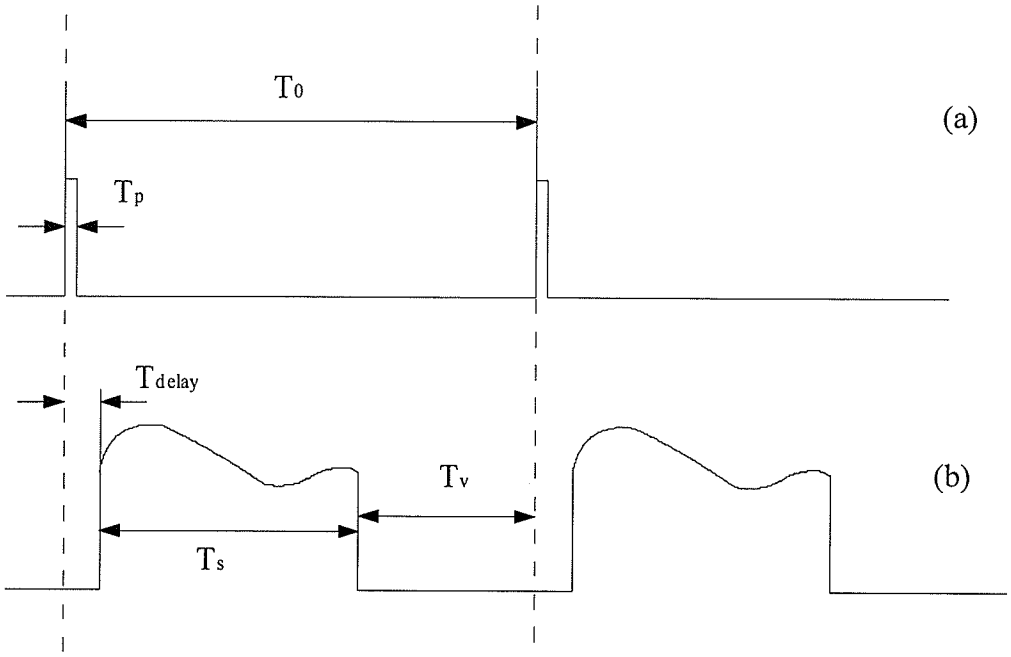


Figure 3.6 Time sequence of CCD array operated under auto-trigger mode (a) trigger signal; (b) CCD array output. T_v is the void scanning time between each scan.

3.1.3 Data acquisition system

The output signal from CCD array detector had a voltage level of 5 V when the detector was saturated by incoming light, and an output voltage of 0.009 V when the sensing elements were totally shielded from light radiation. The CCD array detector itself could scan at 0.25 MHz and a complete scan of the 2048 sensing elements took 8192 μs. Trigger signal generated by the CCD array was used to control the timing of data acquisition. Based on measurement requirements, the NI-6013 (National Instruments, Austin, TX) data acquisition (DAQ) board was chosen (Appendix E). The NI-6013 device provided a very economic solution with high measurement precision and enough sampling rate. The NI-6013 featured 16 channels (eight differential) of 16-bit resolution analog input, a 68-pin connector, and eight lines of digital I/O. The NI-6013 device used the NI data acquisition system timing controller (DAQ-STC) for time-related functions.

The device used the programmable function input (PFI) pins on the I/O connector to connect the device to external circuitry. These connections were designed to enable the NI 6013 to both control and be controlled by other devices and circuits.

The selected accessory set consisted of a 1m long SH68-68-EP shielded cable, and a low-cost, unshielded CB-68LP I/O connector block (National Instruments, Austin, TX). The SH68-68-EP cable was chosen for its very good noise rejecting property. The CB-68LP connector board was a low-cost accessory with 68 screw terminals for easily connecting to 68-pin DAQ devices.

The PCI-6013 DAQ card was configured to acquire voltage signal under differential input mode since differential mode had good common-mode noise rejection. Analog input channels ACH0 and ACH8 were used with ACH0 connected to the positive side of the signal (Appendix E). No custom scaling of the incoming signal was applied and the maximum and minimum values expected to be measured were set to 5 V and 0 V. The code width was calculated according to:

$$CW = \frac{InputRange}{gain \cdot R} \tag{3.7}$$

where CW is the code width (V), $InputRange$ is the signal range (V), $gain$ is the signal gain on DAQ card (no unit), and R is the DAQ card resolution (no unit).

Therefore, with no additional signal gain, an input range of 5 V, and 16 bit resolution, the system code width or one least significant bit (LSB) of the digital value was calculated as 76.29 μV ($5/2^{16}$).

The timing of data acquisition was achieved by using external trigger to control the sampling of CCD output. The PFI0 pin was configured as the external trigger input to

control the timing of multiple times of data acquisition. The DGND pin was used as the ground for the trigger signal (Appendix E).

3.1.4 Virtual instrument programs

LabVIEW (National Instruments, Austin, TX) was a graphical programming language that used icons instead of lines of text to create applications. A LabVIEW program consisted of two parts: a front panel and a block diagram. The input controls and output indicators were placed on the front panel as program interface. Codes controlling the program sequence, logic, and corresponding actions were laid out using graphic function modules. The graphic function modules could be connected together in the block diagram to control the front panel objects.

Two VI programs were developed under LabVIEW. One program was designed to monitor the input voltage continuously and mainly used to adjust the optics (fiber optical probe and illuminator) so as not to saturate the CCD array detector. The second program was designed to sample the CCD output with a specified number of samples (default value was 50), at a corresponding sampling rate (default value 5 kilo-samples/s (KS/s)), and for a specified number of trigger circles (default value 50). The program acquired CCD array signal for 10 ms whenever the external trigger started the acquisition by sending a voltage pulse to PFI0 pin. According to Table 3.2 and Fig. 3.6, it was clear that this time period included the time delay T_{delay}, the sample time T_s and a short period in the void scanning time. Total scanning time before saving the data into the computer took about 0.7 s using default settings (50 trigger circles). After the scanning, data were saved into spreadsheet files. The data were arranged in a matrix where columns represented sample numbers and rows represented wavelength variables. An average

spectrum of all acquired spectra (default value was 50) was calculated and appended into the matrix for saving. The flow charts for the programs are shown in Fig. 3.7.

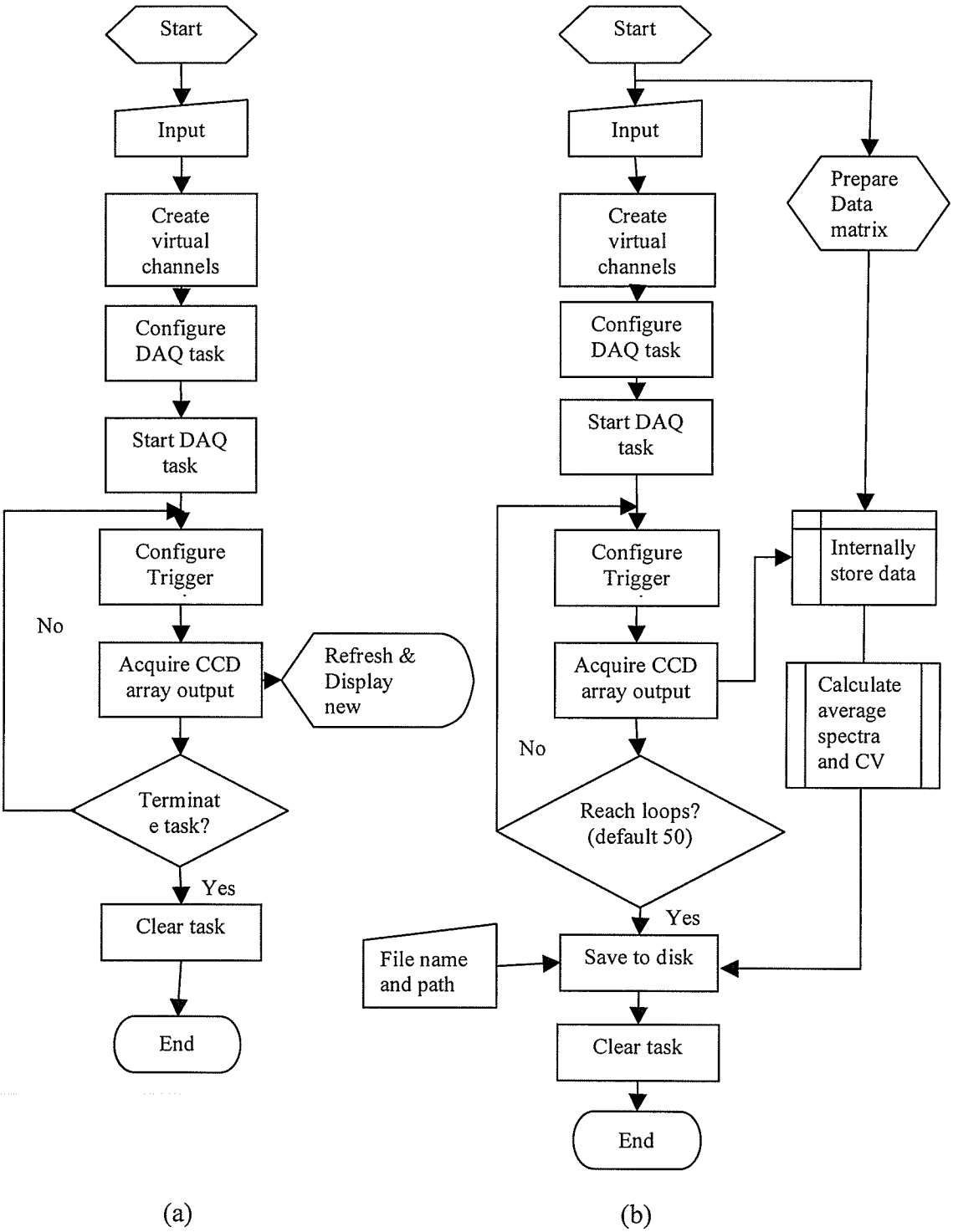


Figure 3.7 Flowcharts of two virtual instrument (VI) programs (a) the monitoring program; (b) the spectra collection program.

3.2 Preparation of Ground Wheat Samples

Hard red spring (HRS) wheat was used in this study. A GAC-2100 Grain Analysis Computer (Dickey-john Corp., Auburn, IL) was used to estimate the initial moisture content of whole wheat. The wheat sample was contained in a fabric bag and placed in a room with constant temperature and moisture level for several days to allow the moisture content of wheat kernel to gain a balance with its environment. Sample to be tested was then taken out from the bag and loaded into the GAC-2100 sample hopper. A moisture content measurement was taken and this procedure was repeated 20 times. The overall moisture average value is 9.16% on a wet basis with a standard deviation of 0.08% and was used as the initial moisture content level.

To prepare wheat samples that could cover a wider range of moisture content (approximately from 5% to 16%), wheat sample of initial moisture content was divided into three portions. One portion of wheat was kept at its initial moisture content level. Moisture content of the second portion of wheat was increased by adding water. The amount of water to be added was calculated according to:

$$W = \frac{K(E - I)}{1 - E} \tag{3.8}$$

where W (g) is the amount of water to be added, K (g) is the initial sample weight, E (%) is the expected moisture content level, and I (%) is the initial moisture content level.

A portion of the desired amount of water was added to ensure that it be absorbed without leftover and the sample was mixed thoroughly by shaking the glass container to allow good exposure of water to all kernels. The container was sealed overnight before the remaining water was added. It was planned to have two moisture content levels above the initial moisture content level (9.16%) with a 3% interval, i.e., approximately 12% and

15%. Since the difference among moisture contents of these samples was relatively small (3%), a large amount of wheat sample (1 kg) was prepared for each moisture content level to ensure that prepared moisture content error can be kept to a minimum. The amount of water to be added was calculated as 34.15 g and 70.72 g. Since it took a while for the wheat kernels to absorb the water thoroughly and the added water to distribute evenly among the kernels, the wheat samples with increased moisture content (1 kg at 12% moisture content and another kilogram at 15% moisture content) were sealed in different glass containers and conditioned over two days at room temperature. Initially, because of the amount of water added, wheat kernels stuck to the wall of glass container and did not peel off easily. The absorption of water was considered complete when no wheat kernel stuck to wall of the glass container. An IB31000P electronic weighing machine (Sartorius AG, Goettingen, Germany) was used to weigh wheat samples. An Adventurer AR1140 analytical balance (Ohaus Corp., Pine Brook, NJ) was used to weigh the amount of water to be added. A Thermolyne OV35025 mechanical oven (Thermolyne, Dubuque, IA, USA) was used to prepare the third portion of wheat by heating the wheat sample in the air oven. Since the relation between heating time and moisture loss was not documented, this procedure was carried out on a trial and error basis. Wheat samples were heated in the air oven at 130°C for an amount of time that was much less than the time required to determine the moisture content. The samples were removed from the oven and placed in the desiccant and allowed to cool down. Then the GAC-2100 Grain Analysis Computer was used to roughly estimate the moisture content level to see if required level of moisture content had been reached. The samples were prepared so as to

prepared so as to have a minimal moisture content of around 5%. After the heating, these samples were kept in desiccant to maintain their moisture levels.

The determination of the reference values of wheat moisture content was carried out using the air oven method. Since the sample that we prepared had a maximal moisture content value of less than 16%, the one-stage procedure was followed. A Cyclotec 1093 sample mill (FOSS Tecator AB, Höganäs, Sweden) was used to grind the sample. A screen with holes of 0.5 mm diameter was used. Several tared aluminum dishes with lids were used to carry the ground wheat samples. The ground wheat samples were heated in the air oven at 130°C for exactly 1 h after the oven attained the desired temperature. Heated wheat samples were placed in the closet with desiccant to allow the samples to cool down before weighing. The moisture was calculated as the weight loss of the ground wheat sample. Detailed procedure can be found in the AACC method 44-15A (AACC, 1983).

For each moisture level, 5 replicates of ground wheat samples were prepared using the lab mill. The Cyclotec 1093 sample mill was dismantled and cleaned before wheat sample of a different moisture content was to be ground. There were 4 moisture content levels and therefore 20 ground wheat samples in total. Spectrum of each ground wheat sample was collected and then the ground wheat sample was transferred to the air oven for moisture measurement according to the standard procedure. Twelve out of the twenty samples were used for calibration development and the remaining eight samples were used as test samples to test the prediction performance of the calibration models.

4. RESULTS AND DISCUSSION

4.1 System Calibration and Testing

An experiment was conducted to calibrate the relationship between the output of spectrograph and wavelength using a single-mode diode laser. The least square estimators of the virtual focal length and central pixel position were calculated using the sequential number of output from CCD detector. The calibrated wavelength relationship was validated using an argon lamp. Then the optical spectral width was experimentally determined. Some of the equations used in system calculation in this section had their origins in the book by James and Sternburg (1969).

4.1.1 Wavelength calibration and validation

After setting up the optical system, the array output from the CCD detector had to be calibrated for corresponding wavelengths. The UL5-3.5G-650C diode laser (World Star Tech Inc, Toronto, Ont.) was used as the monochromatic light source, and an experiment was carried out to acquire the spectra of the diode laser by varying the incident angles. As mentioned in section 3.1.1.5, the relationship of wavelengths versus pixel positions could be determined using Eq. 3.6. To simplify matters, the following calculation was carried out to derive a virtual focal length that could be used to directly get the pixel positions on the CCD array detector. As demonstrated in Fig. 3.4, the physical distance X_n on the array detector could be replaced by pixel positions as represented by the sequential number of the CCD output. In this experiment, a 20 KS/s sampling rate was used and 200 data points were collected. Therefore one sample corresponded to approximately 50 μs and total acquisition took 10 ms sampling period. The chamber of the optical system was

removed temporarily to allow the CCD array detector to be saturated by the ambient light. The sensing element and its range in the acquired data diagram corresponded to those sample points that had an output of about 5 V as shown Fig. 4.1.

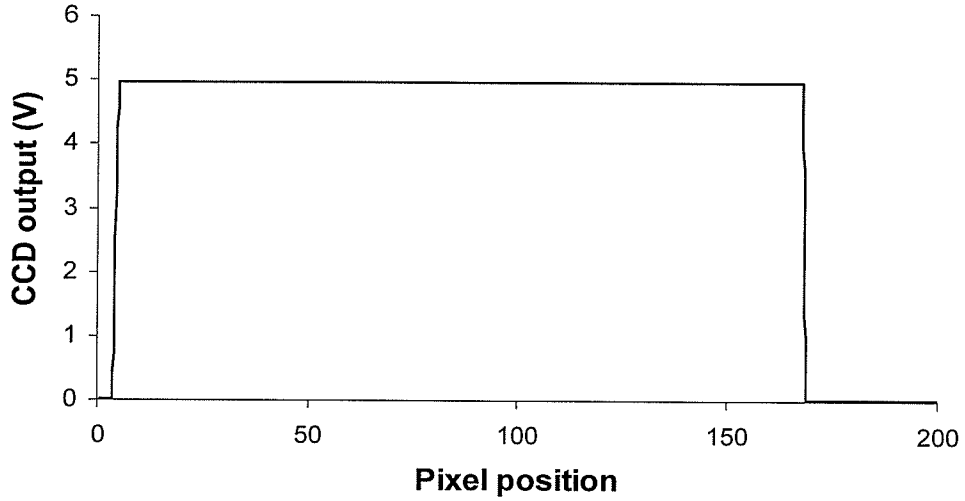


Figure 4.1 Output of the saturated CCD array detector.

From Fig. 4.1, it was evident that the first four samples of CCD array detector at 20 KS/s were near zero (V_{out} = 0.009 V), and sampled voltages rise to 5 V starting from the fifth data point and extend to the 168th data point. The four pixels with zero output at the beginning and 32 zero pixels at the end corresponded to the delay and part of the void scanning time, respectively. The incident angle of the diffraction grating was then set to 0°, and according to the diffraction grating equation (Eq. 2.13), the diffraction angle would be 51.26°. The incident angles were varied and a series of diffraction angles were calculated. Equation 3.6 was used to calculate the pixel position as the incident angles were being scanned instead of wavelengths. Since the optical axis of the collimating path and the imaging path had an angle of $\gamma = 45^{\circ}$, calculation was done to find an incident angle that had a diffraction angle of 45° larger than the incident angle. The incident angle

was 2.47° with a diffraction angle of 47.47° according to the calculation. Several measurements were made with incident angles of -1°, 0°, 2.47°, 3°, 5°. The spectra are shown in Fig. 4.2.

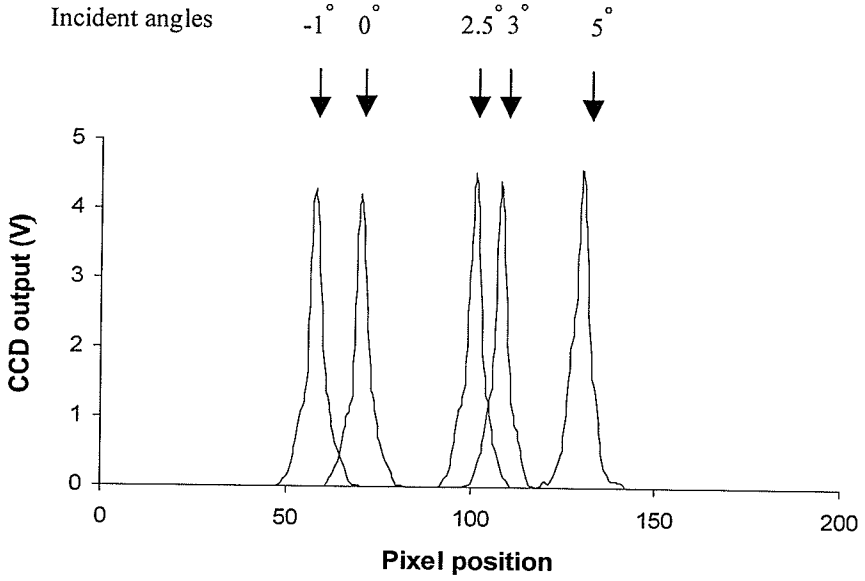


Figure 4.2 Pixel positions on CCD array detector versus incident angles.

To derive the virtual focal length of the imaging lens, the corresponding position of optical axis of the lens should be determined. The virtual focal length was calculated based on experimental data when the incident angles were -1°, 0°, 2.47°, 3°, and 5°. The experimental data are given in Table 4.1.

Table 4.1 Relation of pixels positions versus incident angles

α (degree)	*β (degree)*	*β - α*	*θ (degree)*	Experimental X_n
-1.00	52.89	53.89	-8.89	58.00
0.00	51.26	51.26	-6.26	70.00
2.47	47.47	45.00	0.00	101.00
3.00	46.69	43.69	1.31	108.00
5.00	43.86	38.86	6.14	130.00

α – the incident angle
β – the diffraction angle
θ – the difference between γ (45°) and $\beta - \alpha$.

According to Eq. 3.6, the relationship between pixel position and wavelength was rewritten as:

$$X_n = L_p \times tg(\theta) + X_H \tag{4.1}$$

Therefore, with each incident angle α, a θ value was calculated. The tangential values of θ and corresponding X_n values were regressed against each other using simple linear regression method to estimate values of L_p and X_H. The regression equation is shown in Fig. 4.3.

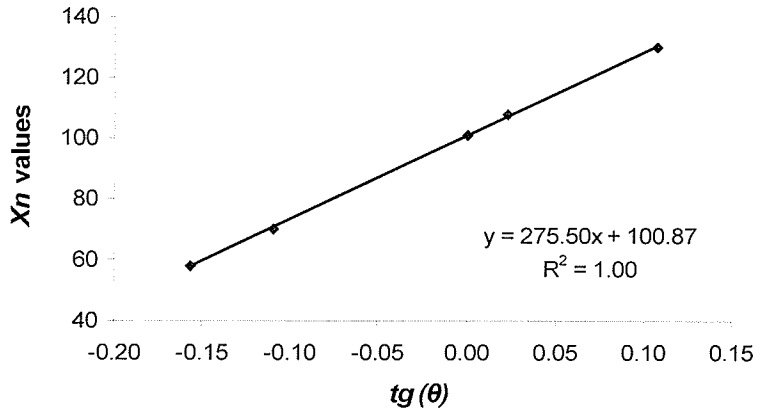


Figure 4.3 Regression equation using experimental values of $tg(\theta)$ and X_n.

Therefore according to the regression equation in Fig. 4.3, the estimated value of L_p was 275.50, and the estimated value of X_H was 100.87. An equation using the two estimated values for parameters L_p and X_H was calibrated and used to determine the wavelengths on the CCD array detector when the system is configured to work as a spectrograph. Maximum and minimum detectable wavelength and spectral range were also calculated. Table 4.2 provides data for incident angles of 0°, 7°, and 17°.

Table 4.2 Maximum and minimum wavelength and spectral ranges versus incident angles

Incident angle (degree)	Maximum wavelength (nm)	Minimum wavelength (nm)	Spectral range (nm)
0	750.18	433.00	317.18
7	890.37	618.10	272.27
17	1067.13	865.90	201.24

The wavelength equation as determined by Eq. 3.6 was validated for the system working under spectrograph configuration using a 6030 Argon lamp (Oriel instruments Stratford, CT). The "Pencil Style" argon lamp source produced narrow, intense lines from the excitation of rare gases. The lamp was connected to a DC supply and operated at a 10 mA input current. The output port of the lamp mount was coupled onto the fiber optical probe to allow only the light from the argon lamp to enter the spectrograph system. Since most of the emissions in the visible to short wavelength NIR region occurred between 640 nm and 840 nm, an incident angle of 7° was selected for the spectrograph which corresponded to a sampling spectral region 618 nm to 890 nm. The emission wavelengths published by the lamp manufacturer and the wavelengths measured by the spectrograph system are listed in Table 4.3. Published emission lines and emission lines acquired using the constructed spectrograph are shown in Fig. 4.4 and 4.5.

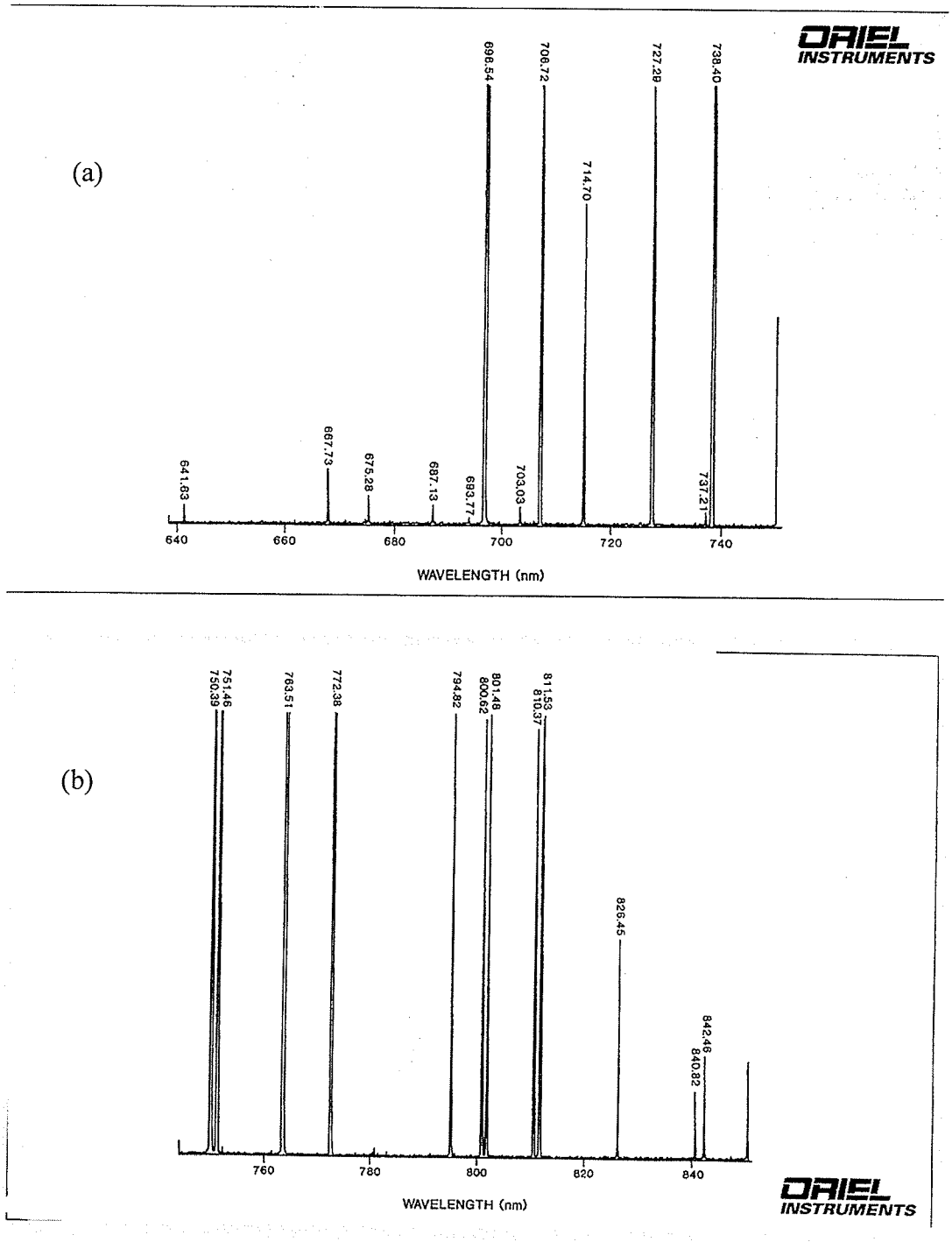


Figure 4.4 Emission spectra provided by Oriel Instruments (a) spectra within the 640 - 740 nm wavelength range; (b) spectra within the 740 - 850 nm wavelength range. *Source:* Oriel Instruments

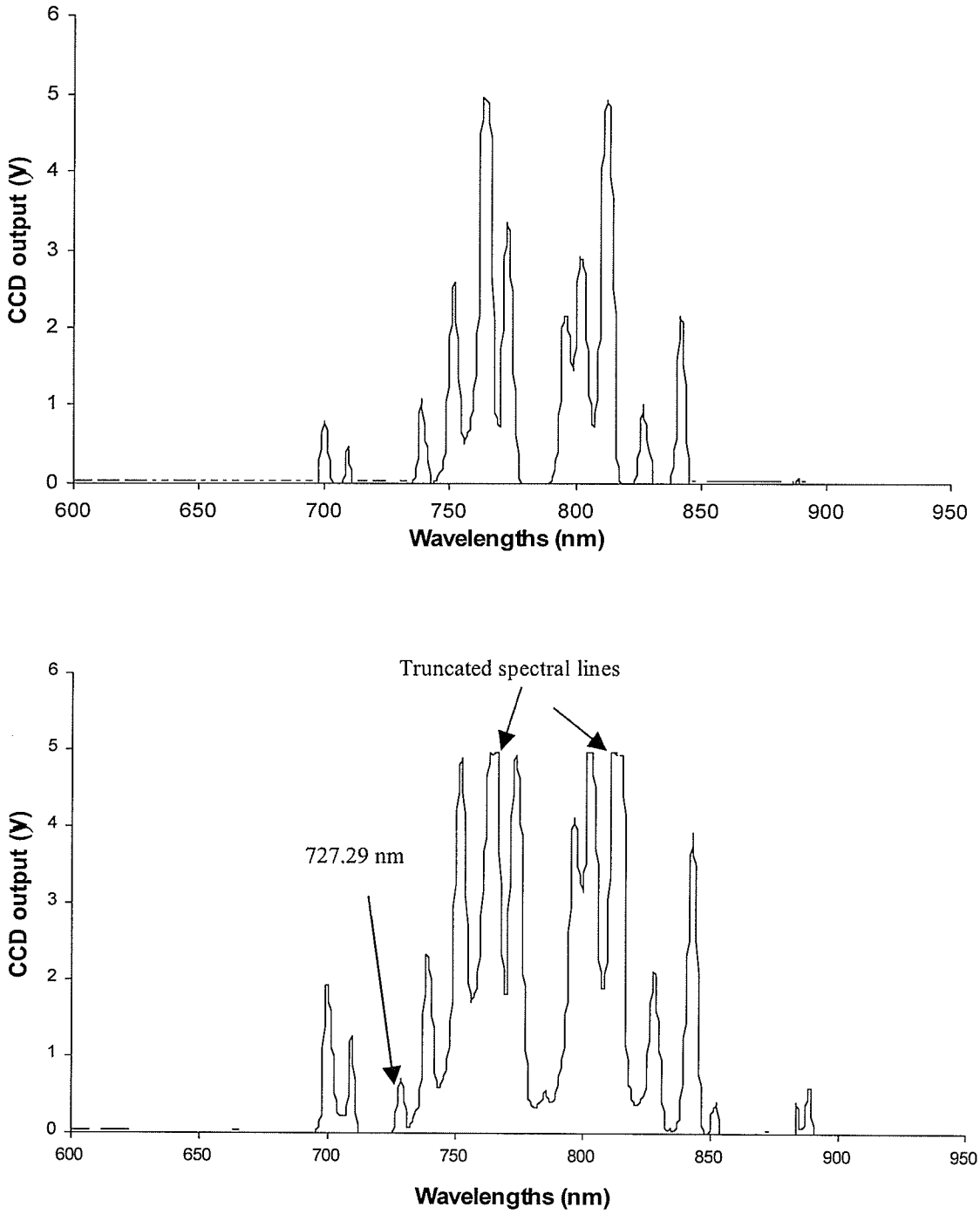


Figure 4.5 Emission spectra measured by the spectrograph (a) without saturation of sensing elements; (b) with several sensing elements saturated by argon lamp input.

Table 4.3 Wavelength validation results using an argon lamp

	Published wavelengths (nm)	Measured wavelengths (nm)	Wavelength difference (nm)	Wavelength error percentage[+] (%)
Discrete wavelengths	696.54	696.48	0.06	0.01
	706.72	706.86	-0.14	-0.02
	727.29	727.12	0.17	0.02
	738.40	737.00	1.4	0.19
	763.51	763.54	-0.03	0.00
	772.38	772.62	-0.24	-0.03
	794.82	796.77	-1.95	-0.25
	826.45	827.87	-1.42	-0.17
Adjacent wavelengths	750.39;751.46	750.47		
	800.62;801.48	801.70		
	810.37;811.53	812.84		
	840.82;842.46	843.16		

[+] Wavelength error percentage is calculated as wavelength difference between published emission wavelengths and measured wavelengths divided by published emission wavelength values.

Table 4.3 summarizes all the emission spectral lines picked up by the spectrograph. It could be seen in the table that the wavelength difference for discrete emission spectral lines were small ($\leq$2 nm) within the entire 272 nm spectral regions. The adjacent wavelengths were where two spectral emission lines were present and the wavelength difference was too small for the spectrograph to resolve. It is obvious from Fig. 4.5 that most of the emission spectral lines were finely resolved by the spectrograph system except those adjacent emission lines with less than 2 nm wavelength differences. Since most emission spectral lines were quite insignificant in spectral spread, the width of these sharp bands corresponded to the line width imposed by the spectrograph system, i.e., the instrumental response. A rough estimate of the line width indicated a FWHM of around 5 nm to 6 nm. A proof to this was the two spectral emission lines at 794.82 nm and 800.62 nm that could just be resolved by the spectrograph system. The limited resolution could also explain the broadening of the spectral lines at the base. It was also

noticed that the relative intensities of the spectral lines appeared to be different compared to Fig. 4.4. The lamp manufacturer's data sheet stated that the phenomenon of varying relative intensities depended on operating conditions. In the spectrograph, the CCD array detector had dead areas between light-sensitive sites and different sensing elements had varying quantum efficiency at different wavelengths. When an emission spectral line incidentally hit upon the dead areas, the emission intensity as represented by the output of the CCD array detector was severely affected. This could be demonstrated by the 727.29 nm emission line. In Fig. 4.5 (a), for all emission spectral lines, care was taken so that the CCD detector was not saturated and the 727.29 nm emission line was not observed. In Fig. 4.5 (b), several CCD detector elements were intentionally saturated using incoming spectral emission which resulted in truncated spectral lines at 763.51 nm and 810.37 nm. The 727.29 nm emission spectral line was observed and the measured 727.12 nm agreed well with the published data. While looking at the published spectra in Fig. 4.4 (a), the 714.70 nm spectral emission line was still missing and this is probably due to both low quantum efficiency and low emission energy of the lamp at this wavelength.

4.1.2 Estimation of optical bandpass

As mentioned in section 3.1.1.5, a slitless spectrograph configuration was selected. Therefore, theoretical optical BP could not be calculated directly. However, after the whole system was set up, the BP as described by FWHM was measured directly using a monochromatic light source. In this experiment, a UL5-3.5G-650C diode laser was used to test the BP. The laser was a continuous single mode laser module with a line width (L) of less than 2 nm at 650 nm. The laser module was also used to adjust the height of

optical post and lens holders, so that the diffracted image could be aligned with the dimension of the CCD array detector.

The incident angle was set to 0° to detect the diffracted light of the laser. The detection wavelength range (433 nm - 750 nm) and the wavelength variables adjacent to the laser central wavelength were calculated using Eq. 3.6. A high acquisition (20 KS/s) speed was used to acquire the spectra. The collected laser spectrum is shown in Fig. 4.6. Assuming the spectral spread of laser output possessed a Gaussian profile, a Gaussian curve was fitted to the measured data (Fig. 4.6). The Gaussian curve had a standard deviation (σ) of 1.4, and the FWHM was calculated as:

$$FWHM = 2\sqrt{-2\log(\frac{1}{2})}\sigma = 2.35\sigma \tag{4.2}$$

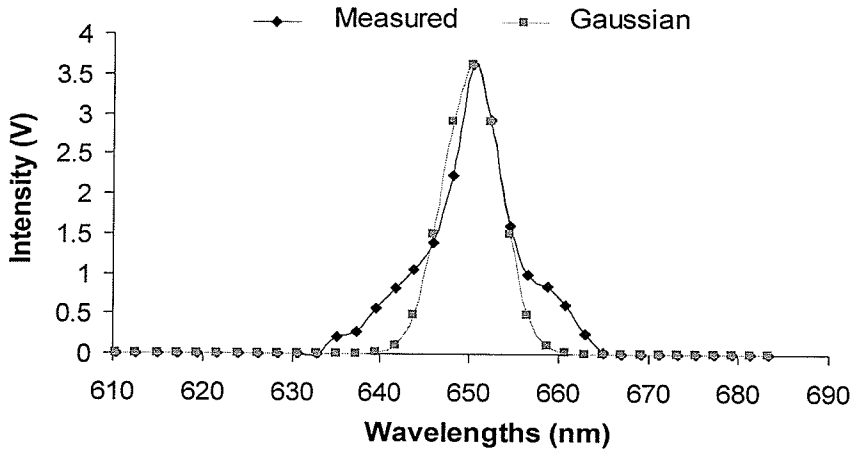


Figure 4.6 Measured single mode laser output and Gaussian fitting curve.

Therefore, FWHM equaled 3.29. Since the central wavelength appeared at the 70th data point, the corresponding wavelengths could be calculated by shifting this value to left and right for 1.645 units (3.29/2). The calculation returned two wavelengths λ_1 and

λ_2 which were 646.9 nm and 652.9 nm respectively. Therefore, the optical BP of the system could be estimated as the difference between λ_1 and λ_2 (i.e. 6 nm).

Assuming the instrumental response was Gaussian and taking the maximum 2 nm line width (L) of the laser diode into account, the ideal optical BP when tested with pure monochromatic light would be:

$$FWHM_A = \sqrt{FWHM^2 - L^2} \tag{4.3}$$

where $FWHM_A$ is the actual FWHM value when tested with pure monochromatic light.

Using $FWHM$ as 6 nm and L as 2 nm, a value of 5.7 nm was obtained for $FWHM_A$.

4.1.3 Estimation of virtual input slit width

Theoretically, optical BP is given by the equation:

$$BP_\alpha = \frac{w\cos\alpha}{knf_A} \tag{4.4}$$

where w is the slit width (mm), α is the incident angle (degree), BP_α is BP corresponding to the incident angle α (nm), k is diffraction order (no unit), n is groove density (grooves/mm), and f_A is the focal length of collimating lens (mm).

When the diffraction grating was used with an incident angle of 17°, the possibly narrowest BP_{17} was 5.5 nm ($FWHM_A \times \cos 17^\circ = 5.70 \times 0.96$). It could also be inferred from the equation the virtual slit width by $w = knf_A BP$. This meant that an entrance slit of 0.26 mm in width seemed to be placed on the focal point of the collimating lens. Thus, without putting extra efforts into fabricating a slit of 0.26 mm in width, a comparable system resolution could be achieved using this slitless spectrograph configuration.

4.1.4 Test of intensity stability

A NI-150 high intensity illuminator (Nikon Inc., Melville, NY) was used as the light source to deliver illumination over the sampling area. The NI-150 illuminator housed a stable high intensity 150 W halogen light source that emitted light continuously in the infrared region. Available fiber optic light guides included two ring lights and double goosenecks. The experimental illumination and spectrograph input light guide setup is illustrated in Fig. 4.7.

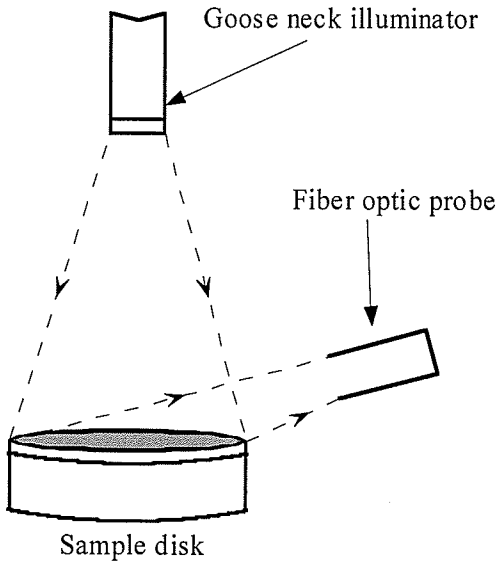


Figure 4.7 Illumination and fiber optic probe setup to collect reflectance spectra. The dashed lines with arrows denote traveling light path.

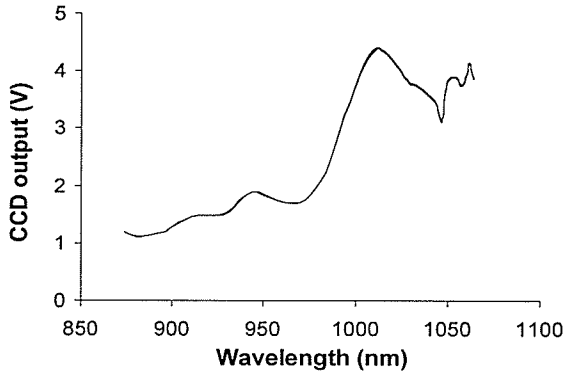
The light spots formed by light delivered using the double gooseneck light guide were arranged to provide even illumination over the sample disk. The fiber optic probe was mounted on a rack to fix the position of the stainless steel end fitting and bending radius of the light guide. To maximize the possibility of collecting diffuse reflectance signal from the sample and avoid effect of specular reflection, the optical axis of fiber optic probe was arranged to be perpendicular to the direction of incoming light from the

gooseneck illuminator. The distance of the double gooseneck light guide to the sample and the intensity of the light were adjusted using real time output graph from the monitor VI program so that reflected light from the standard PTFE disk did not saturate the CCD array detector within the spectral range.

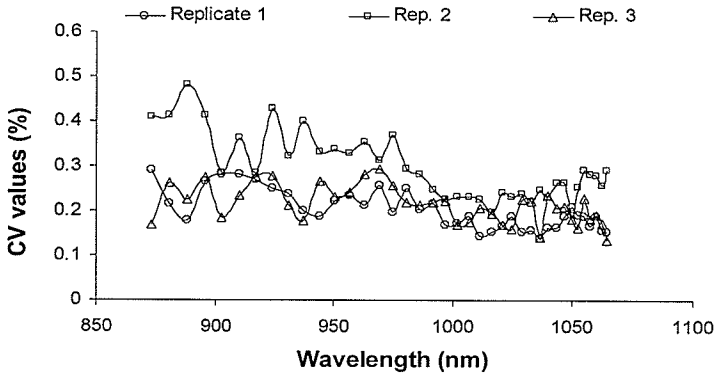
Since the spectrograph had an optical BP of around 5.5 nm under ideal condition, a sampling rate of 5 KS/s was used to acquire the spectra across the entire spectral region which resulted in 50 total sampling data points. Among the 50 sampling data points, 41 points accounted for spectral data. For each spectrum to be saved, 50 sample spectra were acquired and averaged to improve the SNR. An incident angle of 17° was chosen for the spectrograph that resulted in a spectral region from 875 nm to 1064 nm approximately.

Using the spectrograph for quantitatively analysis required the intensity readings be accurate and consistent. An experiment was carried out to examine the intensity stability of the system. Two tests were conducted to find out the intensity reading stability in a short sampling period as well as in a relatively long sampling period. In the first test, 10 reflectance spectra of a standard PTFE disk were collected using the spectrograph every 3 min, which corresponded to a sampling period of 30 min in total. This test was repeated for 3 times. In the second test, 10 reflectance spectra of a standard PTFE disk were acquired using the spectrograph every 3 s, which corresponded to a sampling period of 30 s in total. This test was repeated for 5 times. The relative variability at each sampling wavelength point was measured using coefficient of variation (CV). The 30 reflectance spectra of PTFE disk for the long-term test and CV curves for both short-term and long-term are shown in Fig. 4.8.

(a)



(b)



(c)

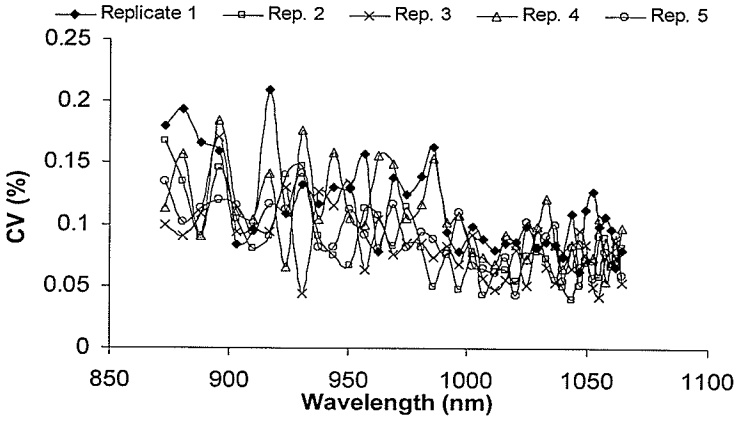


Figure 4.8 Spectra of the PTFE disk and coefficients of variation (CV) curves (a) Thirty reflectance spectra of PTFE disk illuminated by NI-150 illuminator within 90 min; (b) corresponding CV curves for every ten spectra in (a) collected within each 30 min period; (c) CV curves for each ten spectra collected within 30 s.

As is evident in Fig. 4.8(a), the 30 spectra acquired every 3 min within the 90 min sampling period highly overlapped with one another. The overall small values (less than 0.4%) in the CV curve in Fig. 4.8(b) corresponded to this high degree of overlapping of spectra in Fig. 4.8(a). These small percentages of variation indicated that the CCD array detector output in the 30 min period was very stable. The 50 spectra acquired in the short period (30 s) tests were not shown for brevity. In Fig. 4.8(c), the CV values were below 0.2%. This indicated a higher level of intensity reading stability during the short sampling period as compared to the 30 min test. The marginal increase in CV values was inherent in single-beam optical design due to system drift. Therefore, when intensity readings of spectra were important during a long sampling period, the spectrograph should be used with a reflectance standard to correct the drifting effect. Meanwhile, to achieve better performance of the system in moisture measurement experiment, the entire system was powered on for about 30 min before performing measurements so that the whole system could stabilize.

4.2 Spectra of Ground Wheat Collected Using High Resolution Spectrometer

The experiment in the previous section indicated that an optical BP of 5.5 nm could be readily achieved. It was expected that spectral features in the 800 - 1100 nm spectral region should not change drastically across the wavelength dimension that the system could not de-convolve these features. In light of this, a spectrophotometer of higher and adjustable spectral resolution was utilized to acquire spectra of ground wheat in the 800-1100 nm spectral region. The CARY 500 UV-Vis-NIR spectrophotometer (Varian, Palo Alto, CA) was used to acquire the spectra. The spectrophotometer was set to work under diffuse reflectance mode and spectra were recorded as apparent absorbance (log(1/R)). A

Polytetrafluoroethylene (PTFE) disk was used to collect the reference spectra by placing the disk over the sample port. A portion of HRW wheat was ground using the Cyclotec 1093 sample mill. The particle sizes were controlled using a sieve with holes of 0.5 mm diameter. The ground wheat sample was packed into the standard powder cell. The cell had a lid that could be fastened by two screws and this allowed identical packing pressure of wheat powder each time the sample was loaded. In this experiment, the CARY 500 spectrophotometer was set to scan with three different spectral band width (SBW), which were 0.5 nm, 2 nm, and 5 nm, to see if there was any significant spectral difference when the system resolution was reduced. The signal integration time was set to 0.1 s and data were recorded on a 1 nm interval. The spectra of ground wheat acquired with different SBW are shown in Fig. 4.9.

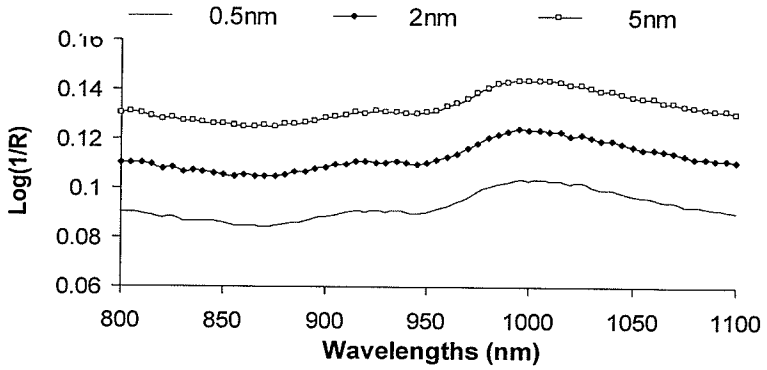


Figure 4.9 Spectra of ground wheat acquired using different spectral bandwidth (SBW) values. The three spectra are intentionally shifted by 0.02 units in log(1/R) values so as to be easily distinguished.

All three spectra in Fig. 4.9 were baseline corrected using the spectra of an identical PTFE disk. The figure clearly illustrates that diffuse reflectance spectra of ground wheat in the 800 - 1100 nm region were smooth and without sharp absorption peaks. Absorption bands were generally broad and featureless.

This feature can be further illustrated by Fig. 4.10. Spectra acquired with a SBW of 0.5 nm was deducted from spectra acquired using SBW of 5 nm and 2 nm. The percentage of such variations that accounts for the total spectral response was calculated. The difference follows random patterns across the whole 800 - 1100 nm region, which indicates that such variations are mainly caused by system noise. There did not exist any spectral feature that could not be resolved using system with a SBW as large as 5 nm. Our system has an optical BP of about 5.5 nm and should not bring about much difference. Therefore, the spectroscopic system could be used to acquire spectra of ground wheat.

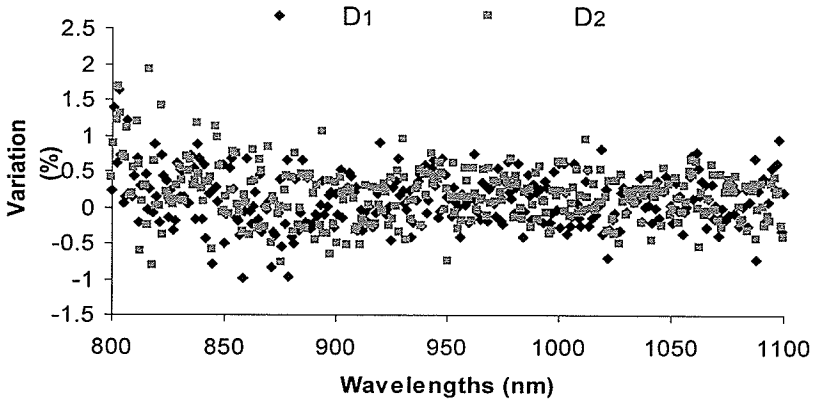


Figure 4.10 Variation percentages for spectra acquired using different spectral bandwidth (SBW) values, D1: the variation percentage between spectra with 2 nm SBW and 0.5 nm, and D2: the variation percentage between spectra with 5 nm SBW and 0.5 nm.

4.3 Determination of Moisture Content in Ground Hard Red Spring Wheat

As part of the study, a preliminary experiment was conducted to demonstrate that the constructed spectrograph system could quantitatively determine the moisture levels in ground wheat at this development stage. It was noted here that the emphasis of this

experiment was not to obtain a prediction model with sufficient accuracy for industrial application.

4.3.1 Spectra collection

In the experiment, a spectrum of the ground whole wheat was acquired 6 times every 3 s and the average spectrum was recorded. The ground wheat samples were prepared according to the procedures described in section 3.2. After a portion of ground wheat was milled, it was transferred from the collection bottle of the mill to a sample tray. A flat spatula was used to scratch the surface of wheat powder in the sample tray to remove extra powder. Then the sample surface was flattened by pressing the flat surface of the spatula against it. Care was taken so that the brim of the sample disk touched the spatula surface to ensure consistent pressure and even surface flatness. The sample disk filled with wheat powder was then put in the center of the illumination spot, and a spectrum scan was performed. To cancel out the effect of the orientation of the sample disk on spectra measurement, another spectrum was taken of the same sample by rotating the sample disk at 90^{o} degrees and average of the spectra was recorded. After scanning the wheat sample, a spectrum of the PTFE disk was taken, which was used as reference spectrum for baseline correction.

As mentioned earlier, the signal was expected to degrade at both ends of the spectral region to be sampled due to vignetting, two data points at each end were discarded. The 12 spectra for calibration are shown in Fig. 4.11. Spectra in Fig. 4.11(a) were transformed using Eq. 2.18 and baseline corrected. The spectral variance was strongly correlated with light scattering rather than the chemical composition of the sample. This is consistent with the findings of previous researchers (Geladi et al. 1985).

The spectra were therefore corrected using MSC, which assumed the scatter for each sample could be corrected to the same level of an "ideal" sample so that corrected spectra were more strongly correlated with chemical information. The spectra in Fig. 4.11(b) are the same 12 calibration spectra after MSC correction. It is evident that the variation was minimized after MSC correction.

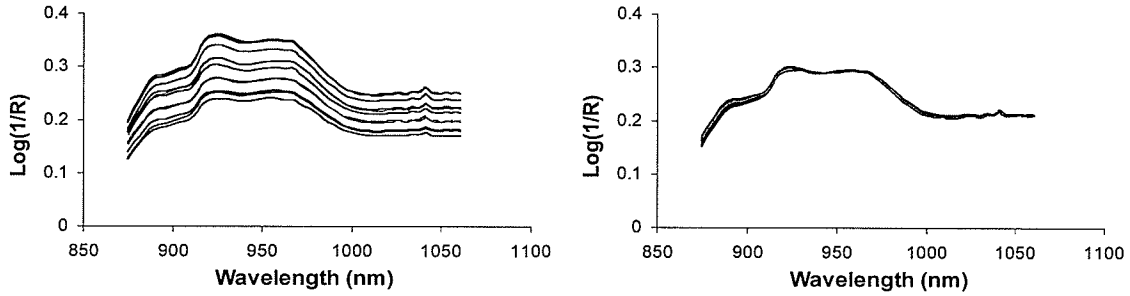


Figure 4.11 Ground wheat spectra for calibration (a) spectra transformed by Log(1/R) and baseline corrected; (b) same group of spectra after MSC correction.

4.3.2 Development of calibration model

Model calibration was done using an add-on software package PLS/IQ of the GRAMS/AI software (Thermo Electron Corp., Waltham, MA). Two data preprocessing methods, MSC and SNV followed by DT, were compared in preparation of the spectral data. The regression method was PLS regression and full cross-validation was used to optimize the model. The diagrams of standard error of cross-validation (SECV) versus model factors are shown in Fig. 4.12.

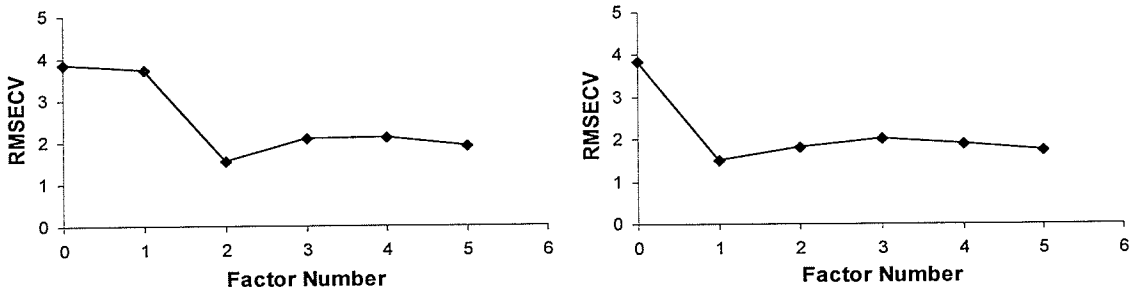


Figure 4.12 Standard errors of cross-validation (SECV) values versus factor numbers (a) calibration spectra corrected with MSC; (b) calibration spectra corrected with SNV followed by DT.

The curves in Fig. 4.12 comply perfectly with theory. The SECV was expected to decrease initially with each addition of factors before reaching a minimum and then went up again as more factors were added. This was true since insufficient number of factors in the model could not account for the constituent of interest while adding factors after the SECV reaching its minimum tended to include factors that correlate with irrelevant noise rather than with constituent of interest. In Fig. 4.12, it is evident that two factors were optimum for model built on spectra corrected with MSC and one factor was sufficient for model built on spectra corrected with SNV followed by DT. The results of cross-validation indicated no sample was to be excluded as an outlier.

The beta coefficients for the two models had similar trends in Fig. 4.13 (Appendix D). The curves (two model factors for MSC corrected spectra and one model factor for SNV corrected spectra) indicated that wavelength variables around 910 nm and 950 nm were important for both models in this experiment. The comparatively large values of beta coefficients for model built on MSC corrected spectra indicated that the model was more sensitive to the changes in these spectral regions.

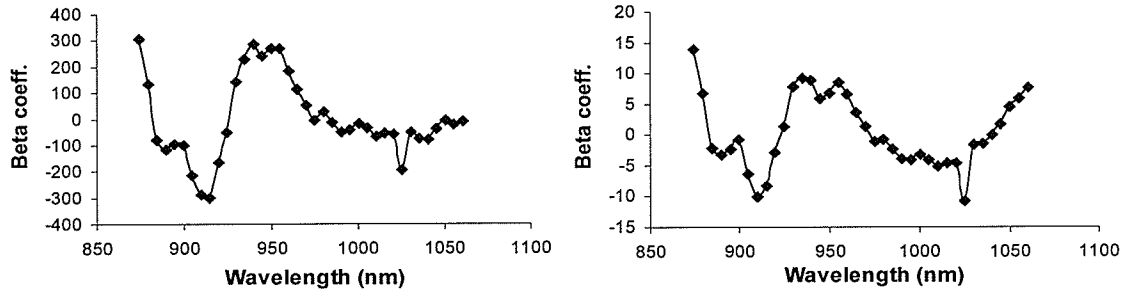


Figure 4.13 Beta coefficients of the calibration models (a) MSC corrected spectra; (b) spectra corrected using SNV followed by DT.

The calibration models were used to predict calibration samples and the calibration regression equations are shown in Fig. 4.14. Calibration and regression statistics are listed in Table 4.4. The statistics in Table 4.4 indicated that preprocessing spectra with SNV followed by DT provided calibration spectra that offered a simpler and more accurate model in this experiment.

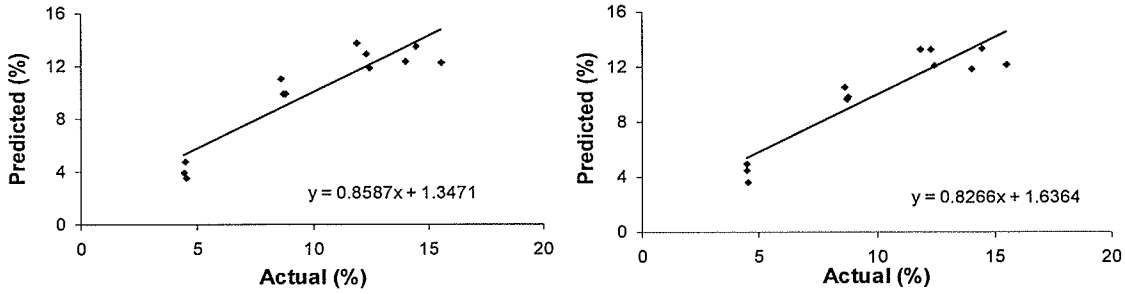


Figure 4.14 Regression equations for calibration models (a) model calibrated with MSC corrected spectra; (b) model calibrated with SNV corrected spectra.

Table 4.4 Model regression statistics

	Factor Number	RMSECV[+]	Slope	Bias	R^2
MSC**	2.00	1.54	0.86	1.35	0.84
SNV and DT***	1.00	1.49	0.83	1.64	0.85

** Model statistics using spectra after MSC correction

*** Model statistics using spectra after SNV and DT correction

[+] RMSECV = root mean squared error of cross-validation values

4.3.3 Validation of model using independent test set

The regression statistics in Table 4.4 only indicated how well the model could describe the variance of the calibration spectra, which actually had been included in the training of the model. To test the prediction ability of calibration models on new samples that had never been used during the calibration development, a separate set of eight spectra collected under same conditions, were used to estimate the prediction error. The regression equations were developed and Working-Hotelling 95 percent confidence bands were constructed for the regression lines. The confidence bands denoted a region where the regression lines lay. The Working-Hotelling confidence bands given in Fig. 4.15 were boundaries used to assess the appropriateness of regression equation with a 95 percent confidence interval. Regression statistics are summarized in Table 4.5.

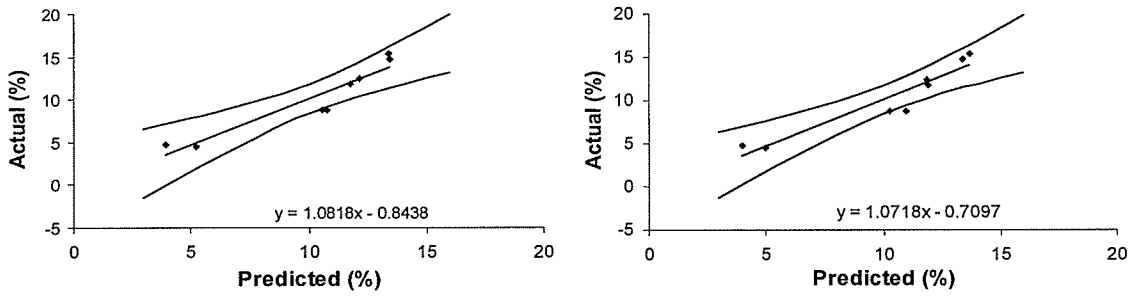


Figure 4.15 Regression equations for validation (a) model calibrated with MSC corrected spectra; (b) model calibrated with SNV corrected spectra.

Table 4.5 Regression statistics of validation

	RMSEP[+]	Slope	Bias	R^2
MSC	1.34	1.08	-0.84	0.88
SNV and DT	1.28	1.07	-0.71	0.89

[+] RMSEP = Root mean square error of prediction.

As can be seen in Fig. 4.15, the 95 percent confidence bands were hyperbolas, which indicated how precisely the regression equation had been estimated. From this

experiment, it was concluded that the regression equation was fairly precisely estimated, which meant that calibration models built in the previous section can be used to determine the moisture levels in ground wheat fairly precisely. The regression equation estimation error was larger at both high and low moisture levels than the median values. Values outside the 95 percent confidence bands corresponded to some moisture levels that could not adequately be predicted by the calibration model, e.g. in Fig. 4.15 (b), a predicted moisture level is 10.96% while the actual moisture level is 8.75%. This might be due to some measurement variation that was so large that it could not be accounted for by the calibration model and resulted in inaccurate prediction. The statistics in Table 4.5 shows that model calibrated with spectra corrected using SNV and DT performs slightly better than model calibrated with spectra corrected using MSC. The actual moisture content and predicted values using calibrated models are listed in Table 4.6.

Table 4.6 Actual moisture values versus predicted moisture content

Sample No.	Actual[+] (%)	MSC[**] (%)	SNV and DT[***] (%)
1	15.36	13.36	13.67
2	14.74	13.44	13.36
3	4.53	5.22	5.00
4	4.78	3.94	3.98
5	8.75	10.77	10.96
6	8.72	10.58	10.26
7	12.42	12.14	11.86
8	11.81	11.77	11.88

[+] Actual moisture content level;
** predicted moisture content using MSC corrected calibration model;
*** predicted moisture content using SNV and DT corrected calibration model.

5. CONCLUSIONS

A spectrograph system based on diffraction grating was designed and built. The spectrograph had tunable scanning wavelength range from the visible-to-near-infrared region with effective wavelength ranging approximately from 450 nm to 1100 nm. Spectral data within 200 nm to 300 nm could be acquired simultaneously at selected sampling wavelength interval. The spectrograph system was comparatively compact in size, robust in structure and relatively fast in spectra collection with good signal-to-noise-ratio. It was integrated with a fiber optic probe to collect light reflectance signals. The optical system of the spectrograph was designed to have a good system throughput with sufficient spectral resolution. The spectrograph used a CCD-2000M charge coupled device (CCD) array detector to acquire the spectra and control the synchronization of data acquisition. The developed software control programs of the system could be used to monitor the spectral input in real time or to record the spectral data and save the data into spreadsheet files. The software was also capable of monitoring the stability of the spectrograph system with a specified time interval and within an arbitrary length of time.

Relation of wavelength and the order of CCD output were calibrated using a single-mode diode 650 nm visible laser and validated with an argon lamp. Most of the wavelengths at eight single emission lines and four adjacent line groups were fairly precisely determined with a maximum wavelength error percentage of 0.26%. An optical BP of 5.7 nm was determined using the 650 nm single-mode diode laser. This experimentally determined value of bandpass (BP) corresponded to a virtual input slit width of 0.26 mm. A stability test indicated that the coefficients of variation (CV) values below 0.16% could be easily achieved across the spectral region of interest.

The performance of the spectrograph was evaluated by estimating moisture content of ground wheat samples using reflectance spectra. Both MSC and SNV followed by DT were used to correct light scattering effect of the reflectance spectra. Calibration model using spectra corrected with SNV followed by DT provided a simpler model. Both models achieved a RMSECV around 1.5% and Coefficient of determination (R^2) around 0.85. An independent test sample set was used to validate the calibrated model. Both calibration models performed well on the test set with root mean squared error of prediction (RMSEP) of around 1.3% and coefficient of determination (R^2) of around 0.88.

6. RECOMMENDATIONS

The developed spectrograph possessed desired characteristics such as good wavelength accuracy, spectral bandwidth, and intensity stability. In the realm of this thesis, it was demonstrated that the spectrograph can be utilized to quantitatively determine moisture levels in ground wheat. Some recommendations to enhance the performance of the system are listed below:

1. It was advised the system be operated under stable surrounding conditions. This was because the lab-scale system, which used test devices instead of integrated design, was subject to influence of harsh environmental conditions such as:

Temperature and moisture level changes: System performance became unstable if there was a violent fluctuation in temperature and relative humidity of environment.

Dust and mechanical vibration: The system was tested within a short distance of a working Cyclotec 1093 sample mill, which degraded its stability.

Instability of the ordinary power source voltage: This was a major factor affecting the stability of intensity reading of the spectrograph.

To reduce the effects of aforementioned influencing factors and make the system more robust towards industrial situations, integrated system design should replace those adjustable mechanical mounts installed in the spectrograph chamber. The integrated design also further reduced the physical dimensions of the system. Illumination light source could also be built into the system with a bifurcated fiber optic probe to make the system portable. A double optical beam scheme made the requirement of stabilized power source less demanding.

2. To enhance results of quantitative analysis of samples under reflectance measurement, several sampling accessories could be added:

The current system tried to maximize the collection of diffuse reflectance light and avoid the effect of specular reflection by setting the optical fiber perpendicular to the illumination light propagation path. An integrating sphere adapted to the fiber optic probe could greatly enhance the diffuse reflectance collection efficiency.

Although great care was been taken, the manually prepared sample disk might have introduced measurement errors due to a difference in the sample surface evenness and sample powder compactness. A sample disk that could ensure consistent sample homogeneity, compactness, and surface smoothness would make the task much easier.

3. To have better spectral specificity and sensitivity in the selected spectral region, there were several ways to improve the system:

In order to perform spectral measurement in the NIR region, an illumination source rich in NIR content, like the halogen-tungsten lamp should be used. The short NIR spectral window in the 850 - 1100 nm was not as rich in chemical information as the region of longer wavelengths. To access longer wavelength spectral region, the grating and CCD array detector should be replaced, and the optical system be adjusted.

For different spectral regions, optical filters with different filtering characteristics could be inserted into either the spectrograph system or the illumination system. The level of stray light and scattered light within the spectrograph system could be brought down by installing optical trapping devices to eliminate unwanted light due to reflection and any other orders of diffraction by the grating.

7. REFERENCES

Allard, F.C. 1990. *Fiber Optics Handbook: for Engineers and Scientists.* New York: McGraw-Hill.

Alphalas GmbH. 2004. *Instructions for using the CCD-2000M array.* Goettingen, Germany: Alphalas GmbH.

American Association of Cereal Chemists. 1983. *Approved Methods of the American Association of Cereal Chemists*. St. Paul, MN: AACC.

Archibald, D.D. and S.E. Kays. 2000. Determination of Total Dietary Fiber of Intact Cereal Food Products by near-infrared Reflectance. *Journal of Agricultural and Food Chemistry* 48: 4477-4486.

Azzouz, T., A. Puigdoménech, M. Aragay and R. Tauler. 2003. Comparison between different data pre-treatment methods in the analysis of forage samples using near-infrared diffuse reflectance spectroscopy and partial least-squares multivariate calibration method. *Analytica Chimica Acta* 484(1): 121–134

Barnes, R.J., M.S. Dhanoa and S.J. Lister. 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy* 43(5): 772-777.

Ben-Gera, I. and K.H. Norris. 1968. Determination of moisture content in soybeans by direct spectrophotometry. *Israel Journal of Agricultural Research* 18: 125-132.

Berger, A.J., Y.Wang and M.S. Feld. 1996. Rapid, noninvasive concentration measurements of aqueous biological analytes by near-infrared Raman spectroscopy. *Applied Optics* 35(1): 209–212.

Birth, G.S. and G.L. Zachariah. 1976. Spectrophotometry of agricultural products. In *Quality Detection in Foods,* ed. J.J. Gaffney, 6-11. St. Joseph, MI: American Society of Agricultural Engineers.

Blanco, M., J. Coello, H. Iturriaga, S. Maspoch and C. de la Pezuela. 1996. Quantitation of the active compound and major excipients in a pharmaceutical formulation by near infrared diffuse-reflectance spectroscopy with fiber optical probe. *Analytica Chimica Acta* 333(1-2): 147–156.

Bokobza, L. 1998. Near infrared spectroscopy. *Journal of Near Infrared Spectroscopy* 6:3–17.

Brennan, J.F. III, Y. Wang, R.R. Dasari and M.S. Feld. 1997. Near-infrared Raman spectrometer systems for human tissue studies. *Applied Spectroscopy* 51: 201–208.

Brereton, R.G. 2003. *Chemometrics : Data Analysis for the Laboratory and Chemical Plant.* Hoboken, NJ: John Wiley & Sons, Inc.

Burns, D.A. and E.W. Ciurczak. 1992. *Handbook of Near-infrared Analysis*. New York: Marcel Dekker Inc.

Campbell, N.A., J.B. Reece, L.G. Mitchell and M.R. Taylor. 2002. *Biology: Concepts and Connections*, 4th Edition. San Francisco: Benjamin Cummings.

Canada Grains Council. 2003. *Canada Grains Council statistical handbook.* Winnipeg, MB: Canadian Grains Council.

Canadian Standards Association. 1991. *Dimensioning and Tolerancing of Technical Drawings.* Rexdale, ON: Canadian Standards Association.

Carter, E.A., P.M. Fredericks and J.S. Church. 1996. Fourier transform infrared photoacoustic spectroscopy of surface-treated wool. *Textile Research Journal* 66(12): 787–794.

Chau, F. 2004. *Chemometrics: From Basics to Wavelet Transform.* Hoboken, NJ: John Wiley & Sons, Inc.

Christensen, C.M., S.B. Miller and J.A. Johnston. 1992. Moisture and its Mmeasurement. In *Storage of Cereal Grains and their Products*, 4th edition. ed. D.B. Sauer, 39-54. St. Paul, MN: American Association of Cereal Chemists, Inc.

Coates, J. 1998. Vibrational spectroscopy: instrumentation for infrared and Raman spectroscopy. *Applied Spectroscopy Reviews* 33(4): 267-425.

Cowe, I.A. and J.W. McNicol. 1985. The use of principal components in the analysis of near infrared spectra. *Applied Spectroscopy* 39: 257-266.

Dal Cason, T. A. 1997. The characterization of some 3,4-methylenedioxycathinone (MDCATH) homologues. *Forensic Science International* 87(1): 9–53.

Davies, T. 1998. The history of near infrared spectroscopic analysis: Past, present and future - "From sleeping technique to the morning star of spectroscopy". *Analusis Magazine* 26(4): M17-M19.

Davis, S. P. 1970. *Diffraction Grating Spectrographs*. New York: Holt, Rinehart, and Winston.

Ding, Q., G. Small and M.A. Arnold. 1999. Evaluation of nonlinear model building strategies for the determination of glucose in biological matrices by near-infrared spectroscopy. *Analytica Chimica Acta* 384(3):333-343.

Dowell, F.E. 2000. Differentiating vitreous and non-vitreous durum wheat kernels by using near-infrared spectroscopy. *Cereal Chemistry* 77(2): 155-158.

Dowell, F.E. and E.B. Maghirang. 2002. Accuracy and feasibility of measuring characteristics of single kernels using near-infrared spectroscopy. In *Proceedings ICC Conference Novel Raw Materials, Technologies, and Products - New Challenge for Quality Control,* 1-6. Budapest, Hungary. May.

Edmund Industrial Optics. 2004. *Optics and Optical Instruments Catalog*. Barrington, NJ: Edmund industrial optics.

Eilert, A.J. 1991. Acousto-optic tunable filter spectrometry: expanding the limits of near-infrared spectroscopic instrumentation. *NIR News* 2(3): 6-9,

Eilert, A.J., W.J. Danley and X. Wang. 1995. Rapid identification of organic contaminants in pretreated waste water using AOTF near-IR spectrometry. *Advanced Instrumentation Control* 50(Pt. 2): 87–95.

Fontaine, J., B. Schirmer and J. Horr. 2002. Near-infrared reflectance spectroscopy (NIRS) enables the fast and accurate prediction of essential amino acid contents. 2. results for wheat, barley, corn, triticale, wheat bran/middlings, rice bran, and sorghum. *Journal of Agricultural and Food Chemistry* 50: 3902-3911.

Franck, P., J.L. Sallerin, H. Schroeder, M.A. Gelot and P. Nabet. 1996. Rapid determination of faecal fat by Fourier-transform infrared analysis (FT-IR) with partial least squaress regression and an attenuated total reflectance accessory. *Clinical Chemistry* 42: 2015–2020.

Gallaher, K.L., F. Baudais and R. Lester 1996. The need for sample conditioning in process streams analyzed with NIR or MID-IR spectrometers. *Advanced Instrumentation Control* 51(Pt. 1): 139–148.

Ge, Z., R. Thompson, S. Cooper, D. Ellison and P. Tway, 1995. Quantitative monitoring of an epoxidation process by Fourier-transform infrared spectroscopy. *Process Control Quality* 7(1): 3–12.

Geladi, P. and B. Kowalski. 1986. Partial least-squares regression: A tutorial. *Analytica Chimica Acta* 185: 1-17.

Geladi, P., D. Macdougall and H. Martens. 1985. Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Applied Spectroscopy* 39(3): 491-500.

Herschel, W. 1800a. Investigation of the powers of the prismatic colours to heat and illuminate objects. *Philosophical Transaction of the Royal Society of London* 90: 255-283.

Herschel, W. 1800b. Experiments on the refrangibility of the invisible rays of the sun. *Philosophical Transaction of the Royal Society of London* 90: 284-292.

Hewavitharana, A.K. and B. Van Brakel. 1997. Fourier transform infrared spectrometric method for the rapid determination of casein in raw milk. *Analyst* 122: 701–704.

Holler, F., D.H. Burns and J.B. Callis. 1989. Direct use of second derivatives in curve-fitting procedures. *Applied Spectroscopy* 43: 877-882.

Hruschka, W.R. and K.H. Norris. 1982. Least squares curve fitting of near infrared spectra predicts protein and moisture content of ground wheat. *Applied Spectroscopy* 38: 317-322.

International Organization for Standardization. 1982. *Technical Drawings*. Geneva: International Organization for Standardization.

Isaksson, T. and T. Næs. 1988. The effect of multiplicative scatter correction and linearity improvement in NIR spectroscopy. *Applied Spectroscopy* 42(7): 1273-1284.

Isaksson, T. and B.R. Kowalski. 1993. Piece-Wise multiplicative scatter correction applied to near-infrared transmittance data from meat products. *Applied Spectroscopy* 47(6): 702-709.

James, J.F. and R.S. Sternburg. 1969. *The Design of Optical Spectrometers.* London, England: Chapman & Hall Ltd.

Johnsen, E. 1997. How to use on-line NIR in the feed and food industries. *Process Control Quality* 9(4): 205–206.

Jones, R.N. 1985. Analytical applications of vibrational spectroscopy-a historical review. In *Chemical, Biological and Industrial Applications of Infrared Spectroscopy*, ed. J. R. Durig, 1-50. Chichester, U.K.: John Wiley and Sons.

Kitagawa, K., S. Hayasaki and Y. Ozaki. 1997. In situ analysis of sizing agents on fiber reinforcements by near-infrared light-fiber optics spectroscopy. *Vibrational Spectroscopy* 15(1): 43–51.

Kubelka, P. and F. Munk. 1931. Ein Beitrag zur Optik der Farbanstriche. *Zeitshrift für Technische Physik* 12:593-604.

Lásztity, R. and A. Salgó. 2002. Quality assurance of cereals– past, present, future. *Periodica Polytechnica-Chemical Engineering* 46(1–2): 5–13.

Law, D.P. and R. Tkachuk. 1977a. Determination of moisture content in wheat by near-infrared diffuse reflectance spectrophotometry. *Cereal Chemistry* 54: 874-881.

Law, D.P. and R. Tkachuk. 1977b. Near-infrared diffuse reflectance spectra of wheat and wheat components. *Cereal Chemistry* 54: 256-265.

Law, D.P. and A.B. Blakeney. 1996. The Kubelka–Munk equation: some practical considerations. *Journal of Near Infrared Spectroscopy* 4:189–193.

Luo, X. 1997. Color image analysis for cereal grain classification. Unpublished Ph.D. thesis. Winnipeg, MB: Department of Biosystems Engineering, University of Manitoba.

Macaloney, G., J.W. Hall, M.J. Rollins, I. Draper, K.B. Anderson, J. Preston, B.G. Thompson and B. McNeil. 1997. The utility and performance of near-infrared spectroscopy in simultaneous monitoring of multiple components in a high cell density recombinant Escherichia coli production process. *Bioprocess Engineering* 17(3): 157–167.

Maghirang, E.B., F.E. Dowell, J.E. Baker and J.E. Throne. 2002. Detecting single wheat kernels containing live or dead insects using near-infrared reflectance spectroscopy. ASAE Paper No. 023067. St. Joseph, MI: ASAE.

Malley, D.F., P.C. Williams and M.P. Stainton. 1996. Rapid measurement of suspended C, N, and P from precambrian shield lakes using near-infrared reflectance spectroscopy. *Water Resources* 30(6): 1325–1332.

Mark, H. 1991. *Principles and Practice of Spectroscopic Calibration.* New York: John Wiley & Sons, Inc.

Martens, H. and T. Næs 1989. *Multivariate Calibration.* Chichester, U.K.: John Wiley & Sons, Inc.

Martens, H., J.P. Nielsen and S.B. Engelsen. 2003. Light scattering and light Absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. *Analytical Chemistry* 75: 394-404.

Matuszewska, A. 1997. The estimation of changes occurring in the coalification process on the ground of the results of infrared spectroscopy and ultimate analysis of the selected coals and macerals. *Polish Journal of Applied Chemisry* 41(1–2): 75–94.

Mayes, D.M. and J.B. Callis. 1989. A photodiode-array-based near-infrared spectrophotometer for the 600-1100 nm wavelength region. *Applied Spectroscopy* 43(1): 27-32.

McQueen, D. H., R.Wilson, A.Kinnunen and J.E. Paaske. 1995. Comparison of two infrared spectroscopic methods for cheese analysis. *Talanta* 42(12): 2007–2015.

McClure, W.F. and A.M.C. Davies. 1988. Faster Fourier transforms in the analysis of near-infrared spectra. In *Analytical applications of spectroscopy*, ed. C.S. Creaser and A.M.C. Davies, 414-436. London, U.K.: Royal Society of Chemistry.

McClure, W.F. 2003. 204 years of near infrared technology: 1800 - 2003. *Journal of Near Infrared Spectroscopy* 11: 487-518.

Miller, C.E. 2001. Chemical Principles of Near-infrared Technology. In *Near-infrared Technology in the Agricultural and Food Industries*, 2nd edition. ed. P.C. Williams and K.H. Norris, 19-37. St. Paul, MN: American Association of Cereal Chemists, Inc.

Murray, I. and P.C. Williams. 1990. Chemical principles of near-infrared technology. In *Near-infrared Technology in the Agricultural and Food Industries*, ed. P.C. Williams and K.H. Norris, 17-34. St. Paul, MN: American Association of Cereal Chemists, Inc.

Norris, K.H. and P.C. Williams. 1984. Optimization of mathematical treatments of raw near-infrared signal in the measurement of protein in hard red spring wheat. I. Influence of particle size. *Cereal Chemistry* 62: 158-165.

Ogilvie, J.F. 1989. lnfrared Spectroscopy of Diatomic Molecules-the First Century. *Chinese Journal of Physics.* 27(4): 281-296.

Osborne, B.G. 2000. Near-infrared Spectroscopy in Food Analysis In *Encyclopedia of Analytical Chemistry: Applications, Theory, and Instrumentation*, ed. Meyers, R. A., Chichester, U.K.: John Wiley & Sons Ltd.

Osborne, B.G., T. Fearn and P.H. Hindle. 1993. *Practical NIR Spectroscopy: with Applications in Food and Beverage Analysis,* 2nd edition. Essex, England: Longman Scientific & Technical; New York: John Wiley and Sons.

Ozpozan, T. and B.S. SchraderKeller. 1997. Monitoring of the polymerization of vinyl acetate by near-IR FT Raman spectroscopy. *Spectrochimica Acta, Part A: Molecular and Biomolecular Spectroscopy* 53A(1): 1–7.

Paliwal, J. 2002. Digital image analysis of grain samples for potential use in grain cleaning. Unpublished Ph.D. thesis. Winnipeg, MB: Department of Biosystems Engineering, University of Manitoba.

Palmer, C. 2002. *Diffraction Grating Handbook,* 5th Edition. Rochester, NY: Richardson grating laboratory.

Pasikatan, M.C., E. Haque., J.L. Steele, C.K. Spillman and G.A. Milliken. 2001. Evaluation of a near-infrared reflectance spectrometer as a granulation sensor for first-break ground wheat: studies with six wheat classes. *Cereal Chemistry* 78(6): 730-736.

Pasquin, C. 2003. Near Infrared Spectroscopy: Fundamentals, Practical Aspects and Analytical Applications. *Journal of Brazilian Chemistry Society* 14(2): 198-219.

Pedersen, D.K., H. Martens, J.P. Nielsen and S.B. Engelsen. 2002. Near-infrared absorption and scattering separated by extended inverted signal correction (EISC): analysis of Near-infrared transmittance spectra of single wheat seeds. *Applied Spectroscopy* 56(9):1206-1214.

Raghavachari, R. 2001. *Near Infrared Applications in Biotechnology*. New York: Marcel Dekker, Inc.

Ren S.X. and L. Gao, 2000. Simultaneous quantitative analysis of overlapping spectrophotometric signals using wavelet multiresolution analysis and partial least squares. *Talanta* 50: 1163-1173.

Savitsky, A. and M.J.E. Golay. 1964. Smoothing and differentiation of Data by simplified least squares procedures. *Analytical Chemistry* 36(8): 1627-1639.

Shaffer, R.E. and G.W. Small. 1996. Genetic algorithms for the optimization of piecewise linear discriminants. *Chemometrics and Intelligent Laboratory Systems* 35: 87–104.

Shao, X.G., H.Gu, J.H. Wu and Y.Y. Shi, 2000. Resolution of the NMR spectrum using wavelet transform, *Applied Spectroscopy* 54:731-738.

Shreve, O.D. 1952. Infrared, ultraviolet, and Raman spectroscopy. *Analytical Chemistry* 24(11): 1692-1699.

Somsen, G.W., I. Jagt, C. Gooijer, N.H. Velthorst, U.A.T. Brinkman and T. Visser. 1996. Identification of herbicides in river water using on-line trace enrichment combined with column liquid chromatography-Fourier-transform infrared spectrometry. *Journal of Chromatography A* 756: 145–157.

Sperline, R.P., J.S. Jeon and S. Raghavan. 1995. FT-IR-ATR analysis of the silicon/ aqueous solution interface using sputtered silicon thin films to access the 1550–1100 cm-1 spectral region. *Applied Spectroscopy* 49(8): 1178–1182.

Stermer, R., Y. Pomeranz and R.J. Mcginty. 1977. Infrared reflectance spectroscopy for estimation of moisture in whole grain. *Cereal Chemistry* 54: 345-351.

Stewart, J.E. 1970. *Infrared spectroscopy: Experimental Methods and Techniques.* New York: Marcel Dekker, Inc.

Strong, J. 1960 The John Hopkins University and diffraction gratings. *Journal of Optical Society of America* 50:1148-1152.

Sun, Z., A. Ibrahim, P.B. Oldham, T.P. Schultz and T.E. Conners, 1997. Rapid Lignin Measurement in Hardwood Pulp Samples by Near-infrared Fourier Transform Raman Spectroscopy. *Journal of Agricultural and Food Chemistry* 45(8): 3088-3091.

Swinehart, D.J. 1972. The Beer-Lambert law. *Journal of Chemical Education* 39(7): 333-335.

Thodberg, H.H. 1996. A review of Bayesian neural networks with an application to near-infrared spectroscopy. *IEEE Transaction of Neural Networks* 7(1): 56-72.

Vazquez de Aldana, B.R., B. Garcia-Criado, A. Garcia-Ciudad and M.E. Perez-Corona. 1996. Non-destructive method for determining ash content in pasture samples: application of near-infrared reflectance spectroscopy. *Communications in Soil Science and Plant Analysis* 27(3–4): 795–802.

Visen, N.S. 2002. Machine vision based grain handling system. Unpublished Ph.D. thesis. Winnipeg, MB: Department of Biosystems Engineering, University of Manitoba.

Wang, D., F.E. Dowell and D.S. Chung. 2001. Assessment of heat-damaged wheat kernels using near-infrared spectroscopy. *Cereal Chemistry* 78(5): 625-628.

Wang, T.S., C.J. Zhu, J.D. Wang, F.M. Xu, Z.O. Chen and Y.H. Luo. 1995. Study on spectral characterization of infrared flare material combustion with remote high-resolution Fourier-transform infrared spectrometry," *Analytica Chimica Acta* 306(2–3): 249–258.

Wesley, I.J., O. Larroque, B.G. Osborne, N. Azudin, H. Allen and J.H. Skerritt. 2001. Measurement of gliadin and glutenin content of flour by NIR spectroscopy. *Journal of Cereal Science* 34: 125-133.

Wetzel, D.L. 1983. Near-infrared reflectance analysis - sleeper among spectroscopic techniques. *Analytical Chemistry* **55**: 1165A-1171A.

Williams, P.C. 1975. Application of near infrared spectroscopy to analysis of cereal grains and oilseeds. *Cereal Chemistry* 52: 561-576.

Williams, P.C. 1999. Grain grading in Canada by Electronics. http://www.grainscanada.gc.ca/Pubs/confpaper/Williams/Egrading/egrading2-e.htm (2005/5/2).

Williams, P.C. and B.N. Thompson. 1978. Influence of whole wheat meal granularity on analysis of HRS wheat for protein and moisture by near infrared reflectance spectroscopy (NIRS). *Cereal Chemistry* 55:1014-1032.

Williamson, J.M., R.J. Bowling, and R.L. McCreery. 1989. Near-infrared Raman-spectroscopy with a 783 nm diode-laser and CCD array detector. *Applied Spectroscopy* 43: 372–375.

Wilson, J. and J.F.B. Hawkes. 1989. *Optoelectronics, An Introduction.* 2nd Edition. New York: Prentice-Hall.

Workman, J.J. 1996. A brief review of near infrared in petroleum product analysis. *Journal of Near Infrared Spectroscopy* 4(1–4): 69–74.

Workman, J.J., P.R. Mobley, B.R. Kowalski and R. Bro. 1996. Review of chemometrics applied to spectroscopy: 1985-95 (Part 1). *Applied Spectroscopy Reviews* 31: 73-124.

Workman, J.J. 1999. Review of Process and Non-invasive Near-infrared and infrared spectroscopy: 1993–1999. *Applied Spectroscopy Reviews* 34(1&2):1–89.

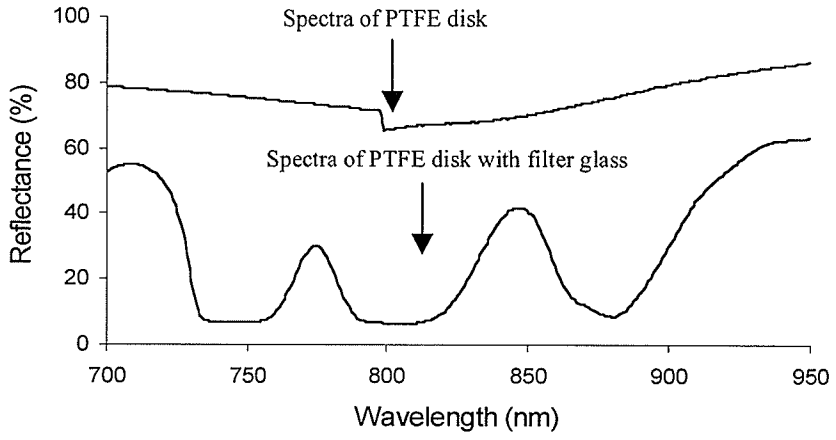
Xie, F., F.E. Dowell and X.S. Sun. 2003. Comparison of near-infrared reflectance spectroscopy and texture analyze for measuring wheat bread changes in storage. *Cereal Chemistry* 80(1): 25-29.

Zhang, L., G.W. Small, A.S. Haka, L.H. Kidder and E.N. Lewis. 2002. Classification of Fourier transform infrared microscopic imaging data of human breast cells by cluster analysis and artificial neural networks. *Applied Spectroscopy* 57(1): 14-22.

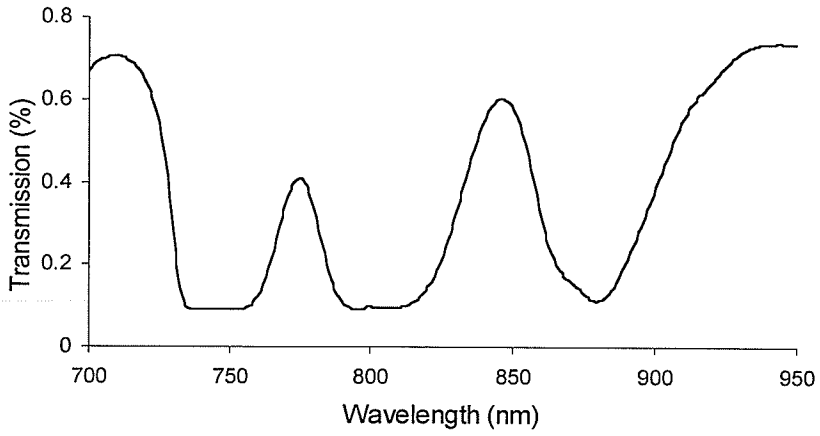
APPENDIX A TESTING OF STANDARD FILTER GLASS

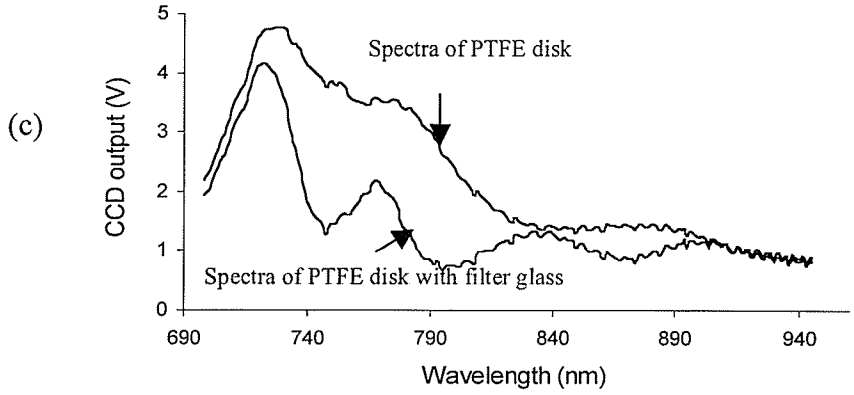
A 5 mm thick piece of BG multi-band filter glass (Esco Products Inc. Oak Ridge, NJ) was used to test the spectrograph system and validate the wavelength and CCD array output relationship. A CARY 500 UV-Vis-NIR spectrophotometer (Varian, Palo Alto, CA) was used for comparison study. A Polytetrafluoroethylene (PTFE) disk was used as the reflectance standard. The PTFE disk had good reflectance in the selected region (700 -950 nm) where three major absorption peaks of the filter glass existed. The original spectra and transmission curves are shown in Fig. A1.

(a)



(b)





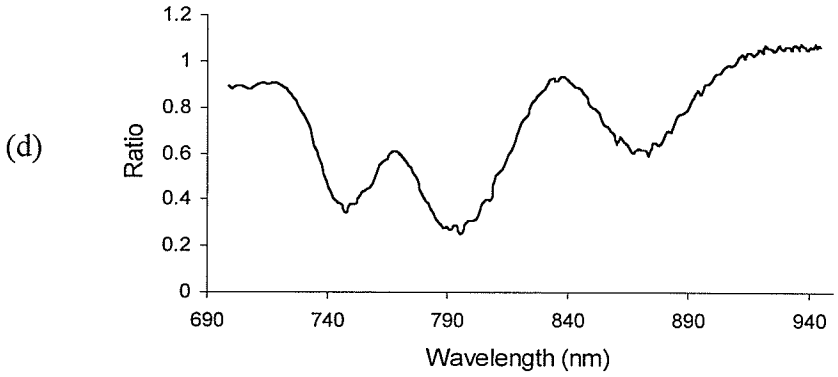


Figure A1 Spectra of reflectance standard with and without filter glass and transmission curves. (a) spectra of PTFE disk measured with and without filter glass on CARY spectrometer; (b) transmission curve derived from spectra in (a); (c) spectra of PTFE disk measured with and without filter glass on the spectrograph; (d) transmission curve derived from spectra in (c)

In Fig. A1(a), the spectra were acquired with a 2 nm spectral band width. The transmission curve in Fig. A1(b) indicated three absorption spectral areas. The three spectral areas center around 750 nm, 800 nm, and 880 nm respectively. The absorption as demonstrated by the spectrograph indicated three absorption areas that center around 750 nm, 796 nm, and 876 nm respectively. The results regarding absorption wavelengths derived from two spectroscopic systems agreed well with each other. A halogen tungsten lamp was used to illuminate the reference standard. The spectra in Fig. A1(c) were

combination of the spectral characteristics of the lamp and the spectrographic system. A high data sampling rate (20KS/s) was used to acquire spectra in this experiment and resulted in a sampling interval ranging from 2 nm on shorter wavelength end to 0.75 nm on longer wavelength end. Therefore the spectra appeared to be stretched out on the longer wavelength end. The spectra and transmission curves in Fig. A1(c) and Fig. A1(d) possessed many spikes. This could be explained by the different quantum efficiency of adjacent sensing element blocks of CCD array detector and readout noise due to electronics.

APPENDIX B: MECHANICAL DRAWINGS OF ANCILLARY PARTS

Note: All drawings have been prepared in AutoCAD 2005 (Autodesk, Inc, San Rafael, CA). Technical presentation principles, dimensioning and tolerancing rules were referred to general principles set by International Organization for Standardization (1982) and standards approved by Canadian Standard Association (1991). All drawings in the appendix were rescaled to fit into the page format and do not necessarily reflect the actual scales of the originals.

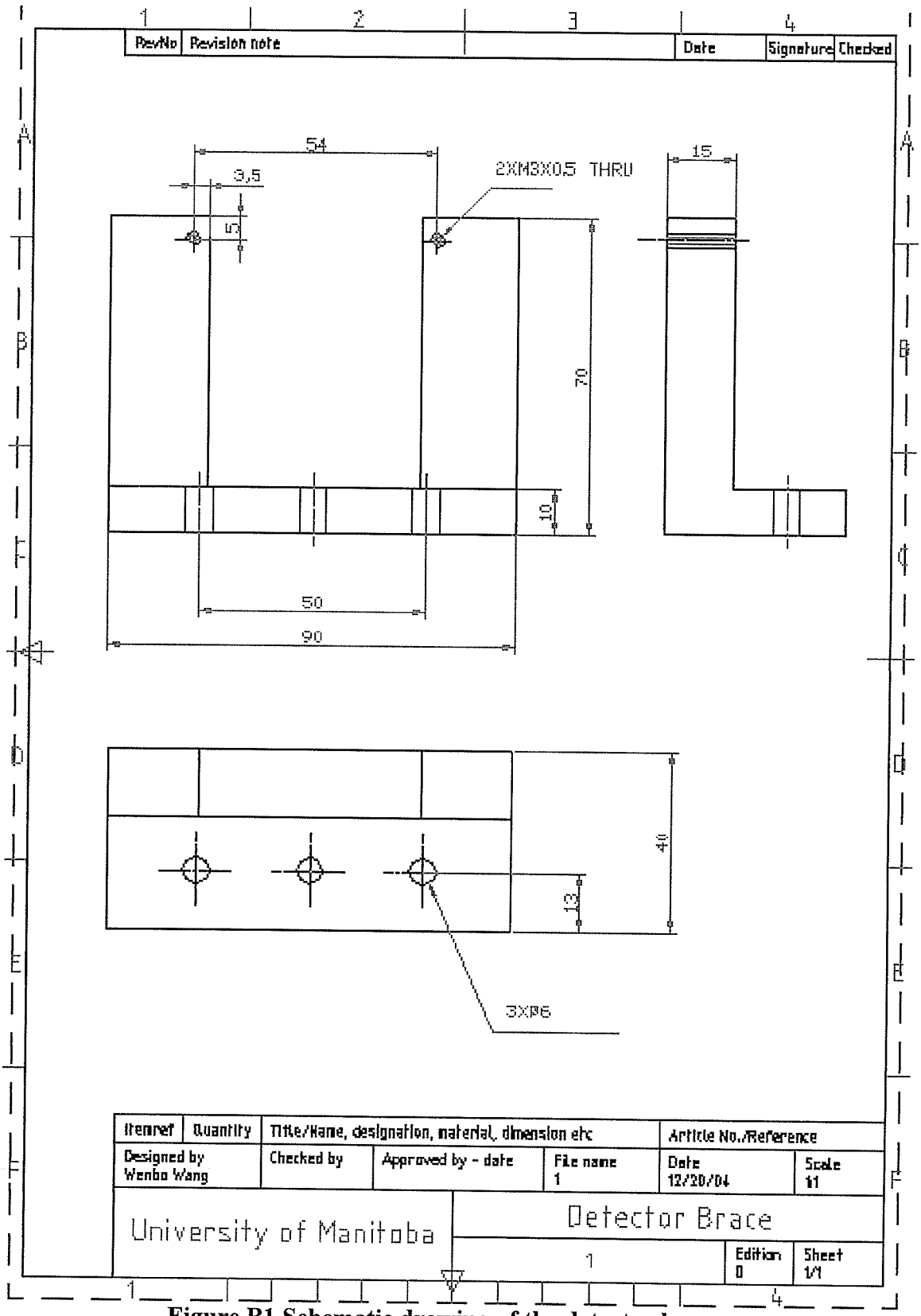


Figure B1 Schematic drawing of the detector brace

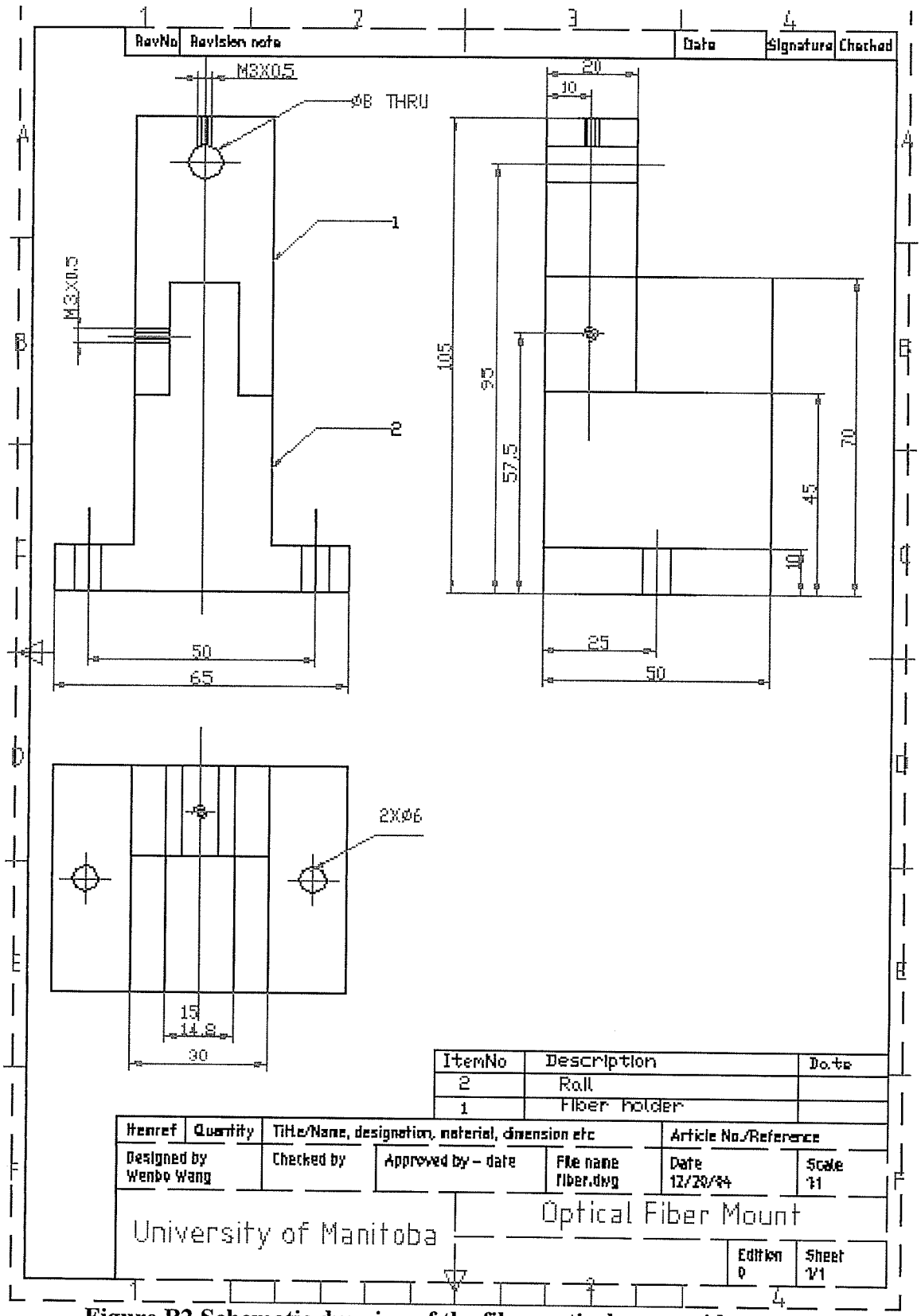


Figure B2 Schematic drawing of the fiber optical waveguide mount

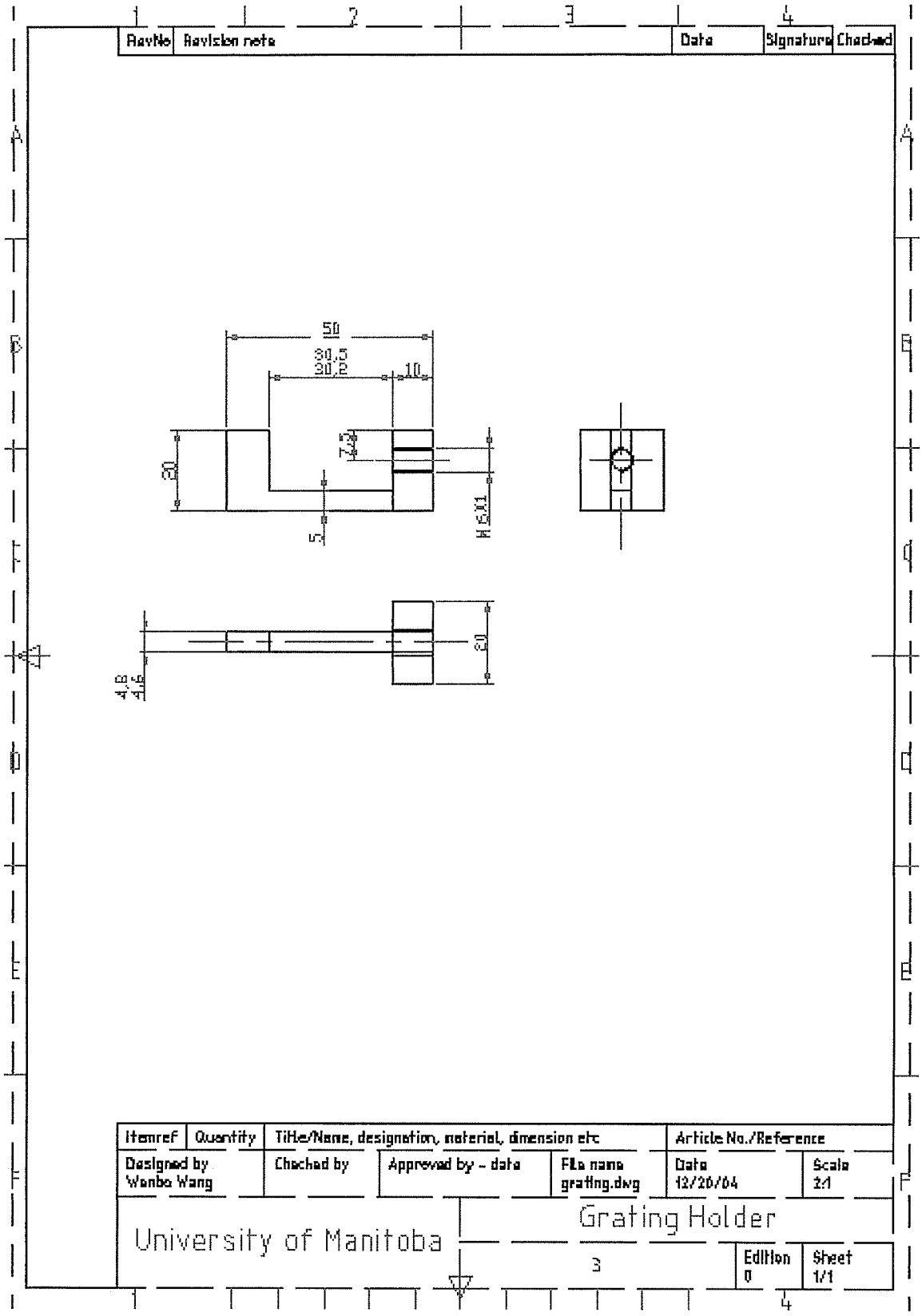


Figure B3 Schematic drawing of the grating holder

APPENDIX C: DIFFRACTION ANGLES FOR THE FIRST POSITIVE ORDER AT AN INCIDENT ANGLE OF 17°

Groove density (Grooves/mm)	Diffracted wavelengths (nm)	α^{+} (degree)	sinα	sinα+sinβ	sinβ	β^{++} (degree)
1200	800	17	0.29	0.96	0.67	41.88
	830	17	0.29	1.00	0.70	44.72
	860	17	0.29	1.03	0.74	47.70
	890	17	0.29	1.07	0.78	50.86
	920	17	0.29	1.10	0.81	54.26
	950	17	0.29	1.14	0.85	57.95
	980	17	0.29	1.18	0.88	62.08
	1010	17	0.29	1.21	0.92	66.87
	1040	17	0.29	1.25	0.96	72.87
	1070	17	0.29	1.28	0.99	82.58

α^{+} = incident angles in degree; β^{++} = diffraction angles in degree.

APPPENDIX D: TECHNICAL ANNOTATIONS

Switched (SW): Under switched mode, output from each individual sensing element of the charge coupled device (CCD) detector are discrete voltage values with respect to each other. There is not voltage output between outputs from each sensing element. Therefore the CCD output consists of evenly spaced output spikes with its number corresponding to the number of sensing elements.

Sample & hold (SH): Under this mode, output from CCD detector is integrated by electronics. The signal from CCD detector is a continuous signal rather than many discrete voltage values in the SW mode.

SubMiniature B (SMB): The SMB is a smaller version of the SubMiniature A (SMA) connector with snap-on coupling.

Transistor-transistor logic (TTL): A common type of digital circuit in which the output is derived from two transistors.

Beta coefficients: Beta coefficients are the regression coefficients calculated using partial least square regression (PLSR). Since PLSR algorithm involves exchanging of score matrices of the spectral data and concentration matrix during matrix decomposition, beta coefficients are dependent on response variables. High peaks and low valleys on beta coefficient curve indicate heavily weighted wavelength variables.

APPENDIX E: NI-6013 DEVICE ACRONYMS AND 68 PIN PINOUT

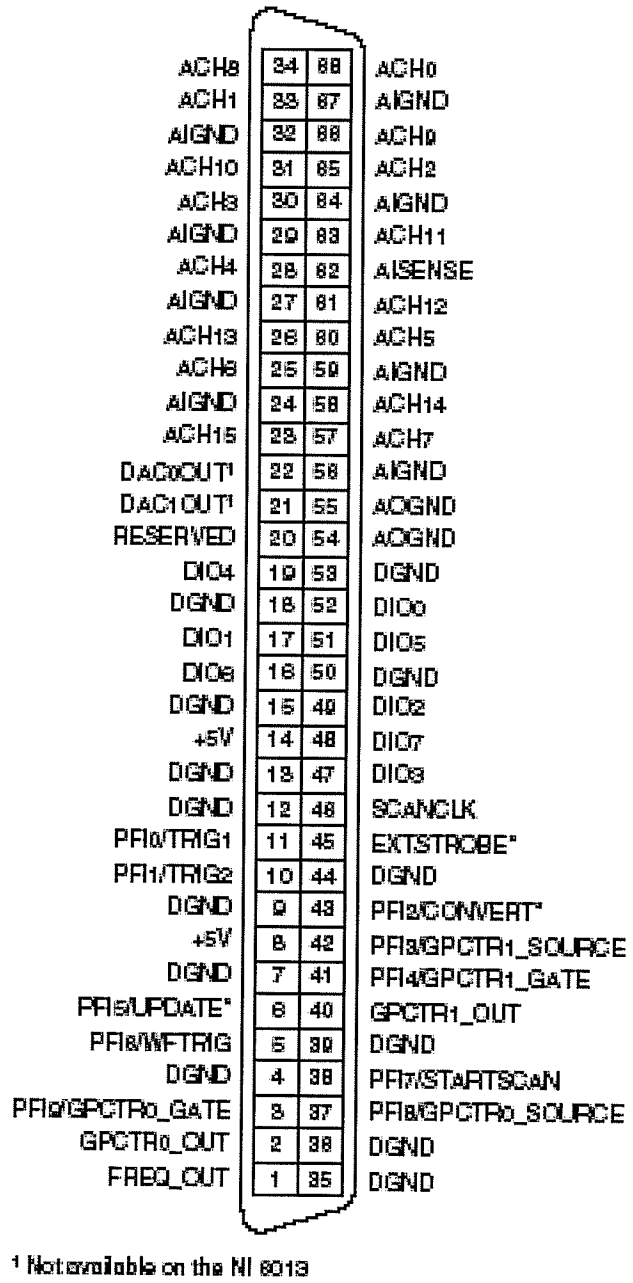


Figure E1 I/O Connector Pin Assignment for the NI 6013/6014. ***Source:*** **National Instruments**

Table E1 Definition of 68 pin MIO I/O connector pinout.

Signal Name	Reference	Direction	Description
PFI0/TRIG1	DGND	Input Output	PFI0/Trigger 1—As an input, this signal is a Programmable PFI. PFI signals are explained in the Connecting Timing Signals section. As an output, this signal is the TRIG1 (AI Start Trigger) signal. In posttriggered DAQ sequences, a low-to-high transition indicates the initiation of the acquisition sequence. In pretriggered applications, a low-to-high transition indicates the initiation of the pretrigger conversions0.
ACH<0..15>	AIGND	Input	Analog Input Channels 0 through 15—Each channel pair, ACH<i, i+8> (i = 0..7), can be configured as either one differential input or two single-ended inputs. ACH0 and ACH8 comprise differential analog input channel 1.
DGND	-	-	Digital Ground—This pin supplies the reference for the digital signals at the I/O connector as well as the +5 VDC supply. All three ground references—AIGND, AOGND, and DGND—are connected together on the device.

Source: DAQ NI 6013/6014 User Manual (National Instruments 2002)

APPENDIX F: MOISTURE RESULTS MEASURED USING AIR OVEN METHOD

Sample No.	Disk weight (g)	Total weight (g)	Total weight after heating (g)	Weight loss (g)	Sample weight (g)	Moisture content (%)
1	17.97	20.83	20.70	0.13	2.86	4.55
2	17.73	19.70	19.60	0.09	1.96	4.78
3	18.19	20.78	20.67	0.12	2.59	4.47
4	18.19	21.09	20.96	0.13	2.91	4.53
5	18.29	20.94	20.82	0.12	2.65	4.50
6	18.11	20.93	20.69	0.24	2.82	8.62
7	20.50	23.29	23.04	0.25	2.79	8.81
8	16.76	18.95	18.76	0.19	2.19	8.72
9	18.17	20.77	20.54	0.23	2.60	8.75
10	18.04	20.85	20.61	0.25	2.81	8.72
11	17.95	20.88	20.53	0.35	2.93	11.86
12	18.46	21.33	20.98	0.35	2.87	12.28
13	14.61	17.23	16.91	0.31	2.62	11.99
14	18.04	20.91	20.57	0.34	2.87	11.81
15	23.17	25.64	25.33	0.31	2.47	12.42
16	18.44	20.77	20.41	0.36	2.34	15.53
17	21.01	23.31	22.96	0.35	2.30	15.36
18	18.15	21.01	20.59	0.42	2.85	14.74
19	18.01	20.21	19.89	0.32	2.19	14.46
20	17.79	20.26	19.91	0.34	2.46	14.00