

Design of a Multispectral Imaging System for Industrial Applications

Project Report - Group G24



Bandara W.G.C. (E/13/036)

Dissanayake D.W.S.C.B. (E/13/078)

Prabhath G.W.K. (E/13/265)

Supervisors: Dr. H.M.V.R. Herath

Dr. G.M.R.I. Godaliyadda

Dr. M.P.B. Ekanayake

Dr. H.A. Suraweera

Department of Electrical and Electronic Engineering

Faculty of Engineering

University of Peradeniya

Sri Lanka

February 2019

Acknowledgements

The success and final outcome of this project required a lot of guidance and assistance from many people and we are extremely privileged to have got this all along the completion of this project. First and foremost, we would like to express our sincere gratitude to our project supervisors Dr. H.M.V.R. Herath, Dr. G. M. R. I. Godaliyadda, Dr. M. P. B. Ekanayake and Dr. H.A. Suraweera for giving us all support and guidance which made us complete the project duly throughout the courses EE 405 - Undergraduate Project I and EE 406 - Undergraduate Project II. We are highly indebted to them for their guidance, inspiration and constant supervision as well as for providing necessary information regarding the project and also for their support in completing the project.

We owe our deep gratitude to Prof. J.B. Ekanayake and Dr. S.G. Abeyratne, the course coordinators for the encouragement, cooperation extended and especially for structuring the course in an effective manner.

We respect and thank to our project guide Prof. T. Madhujith from Department of Food Science and Technology, University of Peradeniya, who took keen interest on our project work and guided us all along, till the completion of our project work by providing all the necessary information for developing a good system. Also we would like to thank Ms. S.M.D. Demini, Instructor of Department of Food Science and Technology, all the staff members of the department and manager of the university farm for the given support.

We would not forget to remember all staff members of Engineering Workshop, Faculty of Engineering, University of Peradeniya, specially the workshop Engineer Mr. M.M.K. Sirisena and workshop director for their encouragement and more over for their timely support and guidance.

Also we would like to thank Dr. Aruna Gunawaradane and members of the circuit design laboratory for given support in making of the circuits.

Also special thanks goes to department of chemical and process engineering for proving lab equipments.

Moreover, we are immensely grateful to the University of Peradeniya research grant [grant number - URG/2017/26/E] and Peradeniya Alumni Association, Melbourne, Australia for all the financial support.

Finally, we wish to extend our sincere gratitude to the lecturers, fellow colleagues and all other individuals who gave their ideas, support and encouragement that empowered us to make our project a success.

Bandara W.G.C. (E/13/036)

Dissanayake D.W.S.C.B. (E/13/078)

Prabhath G.W.K. (E/13/265)

Department of Electrical and Electronic Engineering,
Faculty of Engineering,
University of Peradeniya.

Abstract

Multispectral imaging is a novel technology for obtaining both spatial and spectral information from an object. In recent years, multispectral imaging has rapidly emerged as a powerful and fastest-growing non-destructive tool for food quality analysis and control. Using the multispectral imaging, the spectrum associated with each pixel in a food can be used as a signature to characterize the biochemical composition of each pixel. As a result, multispectral imaging provides the potential for more accurate and detailed information extraction than is possible with other type of technology for the food industry.

As the first part of this research, a multispectral imaging system was developed based on the selected wavelengths from the Ultra-Violet (UV), visible and Near Infra-Red (NIR) regions of the electromagnetic spectrum with a resolution of nine spectral bands (405 nm, 430 nm, 505 nm, 590 nm, 660 nm, 740 nm, 850 nm and 890 nm). Next, a mathematical algorithm was developed to evaluate the quality of chicken meat stored at 4°C using the Principal Component Analysis (PCA). In the latter part of the research, the developed multispectral imaging system was used to estimate the adulteration level of turmeric powder samples available in the market with the aid of PCA and multivariate statistical analysis methods.

Table of contents

List of figures	viii
List of tables	xii
Nomenclature	xiii
1 Introduction	1
1.1 Multispectral Imaging	1
1.2 Applications of Multispectral Imaging	2
1.3 Terminology	3
1.3.1 Spectral Range	3
1.3.2 Spectral Resolution	3
1.3.3 Spatial Resolution	3
1.3.4 Band Numbers	3
1.3.5 Spectral Signature	4
1.4 Representation of Multispectral Image Data	5
1.4.1 The Data-cube	5
1.4.2 Features of Multispectral Data - "data-cube"	7
2 Statistical Tools Used for Analysis of Multispectral Data	8
2.1 Introduction	8
2.2 Statistical Tools used for Feature Extraction and Dimensionality Reduction	8
2.2.1 Principal Component Analysis (PCA)	9
2.2.2 Fisher's Discriminant Analysis (FDA)	11
2.2.3 Feature Space Discriminant Analysis (FSDA)	14
2.3 Tools used for Clustering	16
2.3.1 Spectral Clustering	16
2.4 Tools used for Classification	17
2.4.1 K-Means	17

3 Analysis of Multispectral Images Taken from the Database	20
3.1 About the Database	20
3.2 Analysis of Multispectral Images	22
3.2.1 Selection of pixels from fake and real strawberries	22
3.2.2 Spectral Signatures of all the Pixels	22
3.2.3 Applying Principal Component Analysis (PCA)	22
3.2.4 Scatter Plot of All the Pixels in Reduced Dimensional Space	24
3.2.5 Applying Spectral Clustering Algorithm	24
3.2.6 Classification of real and fake lemon slices using K-means Algorithm	26
4 Hardware Design of the Multispectral Imaging System	27
4.1 Selection of Spectral Bands	27
4.2 Capturing RGB Images of Each Wavelength Band	29
4.3 Generation of Multispectral Image	32
4.4 Conversion to Reflectance Image	32
5 Meat Quality Assessment Using The Multispectral Imaging System	35
5.1 Introduction	36
5.1.1 Advantages of Using Multispectral Imaging Technology in Food Analysis	36
5.2 Methodology	37
5.2.1 Measurement of Meat Quality Parameters	37
5.3 Algorithm and Results	43
5.4 Conclusion	46
6 Validation of multispectral imaging for the detection of adulterants in turmeric samples	48
6.1 Introduction	48
6.2 Materials and Methods	50
6.2.1 Preparation of samples	50
6.2.2 Multispectral Image acquisition	51
6.3 Multispectral Data Analysis	52
6.3.1 Image Pre-Processing	52
6.3.2 Spectral Signatures	54
6.3.3 Principal Component Analysis (PCA)	56
6.3.4 Bhattacharyya distance and Jeffries-Matusita (JM) separability measure	58
6.3.5 Maximum Likelihood Estimation (MLE)	59

Table of contents	vii
6.4 Results and Discussion	61
6.4.1 Selection of the optimal filter and filter parameters	61
6.4.2 Prediction of Adulteration Level	61
6.5 Conclusion	68
References	69
Appendix A LED Driver Circuit - Schematic and Layout	75
Appendix B Discovery Board Support Circuit - Schematic and Layout	77

List of figures

1.1	Light spectrum [1]	1
1.2	Spectral response of an RGB camera[2]	2
1.3	Idea of spectral signature[3]	4
1.4	The schematic diagram of multispectral image (data-cube) for a piece of meat showing the relationship between spectral and spatial dimensions. Every pixel in the multispectral image is represented by an individual spectrum containing information about chemical composition at this pixel[4].	6
1.5	Unfolding multispectral data "data-cube"	7
2.1	Geometrical illustration of PCA as an orthogonal projection to a lower-dimensional subspace. Points in a plane are projected onto a line.	9
2.2	(a): Samples from two classes projection onto the line joining the class means. Note that there is considerable class overlap in the projected space. (b): Corresponding projection based on the Fisher linear discriminant (FDA), showing the greatly improved class separation.	12
3.1	Multispectral image of a fake and real lemon slices taken from the database[5] for our analysis	21
3.2	Selected pixels from fake and real strawberries (121 pixels from fake strawberries and 121 pixels from real strawberries.)	23
3.3	Spectral signatures of all the selected pixels from fake and real strawberries - <i>Red</i> lines for fake strawberries and <i>blue</i> lines for real strawberries. The x-axis is wavelength(nm) and y axis is intensity value	23
3.4	Spectral signatures of fake and real strawberries in reduced dimensional space (for first ten principal components) - <i>Red</i> lines for fake strawberries and <i>blue</i> lines for real strawberries.	24

3.5	Scatter plot of fake and real strawberries in reduced dimensional space (PC1, PC2 and PC3) - <i>Red</i> circles for fake strawberries and <i>blue</i> circles for real strawberries.	25
3.6	Variation of eigenvalue difference for different sigma(σ) values	25
3.7	Scatter plot of fake and real strawberries after applying spectral clustering algorithm to data from PCA - <i>Red</i> circles for fake strawberries and <i>blue</i> circles for real strawberries.	26
4.1	Power spectral distribution for the LEDs on multispectral imaging device. The color of the curve corresponds to the color of the dominant wavelength. Black and brown colors were used for wavelengths at near infra-red region(NIR)	28
4.2	Integrating hemisphere and LED driver circuit	29
4.3	Implemented Multispectral Imaging System as a Block Diagram	30
4.4	All the parts of the multispectral imaging system - the integrating hemisphere, smartphone camera and discovery board (STM32F0)	30
4.5	The complete procedure of acquiring a multispectral image by Windows batch script as a flow chart	31
4.6	Sample multi-spectral image of a chicken meat sample captured using implemented multispectral imaging system	33
4.7	BLACK and WHITE reference plates used to convert intensity image into reflectance image	34
5.1	Flowchart of the key steps involved in meat quality analysis using the multi-spectral imaging system	38
5.2	The key steps involved with measuring microbial count (CFU/g) and pH value in a lab environment	40
5.3	The method of determining microbial count (Colony Forming Units (CFU) /g with the aid of dilution series.	41
5.4	Variation of microbial count (Colony Forming Units per gram - CFU/g) for storage time of 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs.	41
5.5	Spectral signatures of all the pixels for storage time of 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs. Microbial count for each storage time also indicated in the legend.	43
5.6	Average spectral signatures of chicken meat samples for storage time of 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs. Microbial count for each storage time also indicated in the legend.	44

5.7	Average signature curves for storage time of 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs after performing the Principal Component Analysis (PCA) on the original data set. Now, x-axis represents the six principal components and y axis represents the average score for each storage time.	45
5.8	Average value of first principal component (PC1) for storage time of 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs.	46
5.9	Scatter plot of all the pixels in reduced dimensional space - PC1, PC2 and PC3. In this case pixels are clustered together according to the storage time.	47
5.10	The multispectral imaging system installed in a food processing line to automate quality assessment	47
6.1	Schematic of the developed multispectral imaging system	51
6.2	The power spectral density distribution curves of the LEDs used for LED switching circuit	55
6.3	A multispectral image of a turmeric powder sample and (b) its spectral signature	55
6.4	A Two-dimensional data matrix corresponding to an adulteration level. There are altogether nine data matrixes (for the calibration data set) for the adulteration levels of 0%, 5%, 10%, 15%, 20%, 25%, 30%, 35% and 40%	56
6.5	(a) Variation of mean spectral signature curves at different adulteration levels of turmeric powder (b) Variation of mean intensity with adulteration level for different wavelengths	57
6.6	Complete algorithm as a schematic	60
6.7	The Curve fitted results (variation of mean Bhattacharyya Distance with adulteration level) for the Wiener filtering with window size of 3 pixels \times 3 pixels (blue and pink lines) and for the median filtering with window size of 3 pixels \times 3 pixels (red and black lines)	62
6.8	The Curve fitted results for training data (a) variation of the Bhattacharya Distance with adulteration level (b) variation of the JM Distance with adulteration level	63
6.9	The variation of Bhattacharyya Distance with adulteration level for validation data on the calibrated curve ($B = 0.0010p^2 + 0.0208p$). The correlation coefficient between calibrated curve and the validation data is 0.9644	63
6.10	The Bhattacharyya distance of commercially available samples CS_1 , CS_2 , CS_3 , CS_4 , CS_5 , CS_6 , CS_7 , CS_8 , CS_9 and CS_{10} on the calibration curve	65

6.11 Microscopic views of (a) authentic turmeric, (b) rice flour, and (c) – (i) commercially available turmeric samples $CS_1, CS_2, CS_3, CS_4, CS_5, CS_6, CS_7, CS_8, CS_9$ and CS_{10} respectively. Magnification: 10×10 .	66
6.12 Variability in normalized log MLE score over the surface of turmeric powder samples (a) Sample CS_1 (b) Sample CS_2 (c) Sample CS_3 (d) Sample CS_4	67
A.1 Layout of the LED driver circuit	75
A.2 Schematic diagram of the LED driver circuit	76
B.1 Schematic diagram of the discovery board support	77
B.2 Layout of the discovery board support	78

List of tables

3.1	Details of multispectral images database[5]	22
3.2	Results of classifying fake and real strawberries using k-means algorithm after performing PCA and spectral clustering	26
4.1	Details of Narrow-band LEDs used in the multispectral imaging system . .	28
5.1	Meat Quality Parameters for Different Storage Times at 4 °C	42
6.1	Details of the LEDs used in the LED switching circuit (Purchased from Digi-Key Electronics - https://www.digikey.com)	52
6.2	The performance of curve fitting results under the Wiener filtering (window size: 3 pixels × 3 pixels) and the median filtering (window size: 3 pixels × 3 pixels)	61
6.3	Details of estimated adulteration levels of commercial samples	64

Nomenclature

Acronyms / Abbreviations

CCD Charge-Coupled Device

CFU Colony Forming Units

CIE International Commission on Illumination

FDA Fisher's Discriminant Analysis

FSDA Feature Space Discriminant Analysis

HSI Hyperspectral Imaging

LDA Linear Discriminant Analysis

MSI Multispectral Imaging

NIR Near Infra-red

PCA Principal Component Analysis

RGB Red-Green-Blue

TPC Total Plate Count

UART Universal Asynchronous Receiver-Transmitter

UV Ultra-Violet

Chapter 1

Introduction

1.1 Multispectral Imaging

The human visual system senses the electromagnetic spectrum from approximately 400 nm to 700 nm with sensitivity in three overlapping spectral bands that make up the colors we know Red, Green and Blue (RGB) as shown in Fig. 1.1[1, 6].

Conventional cameras have a similar response to the human visual system by design[2] (see Fig. 1.2), as they are typically used to capture images for human presentation and consumption. Due to the fairly broad absorption spectrum of the color response of the

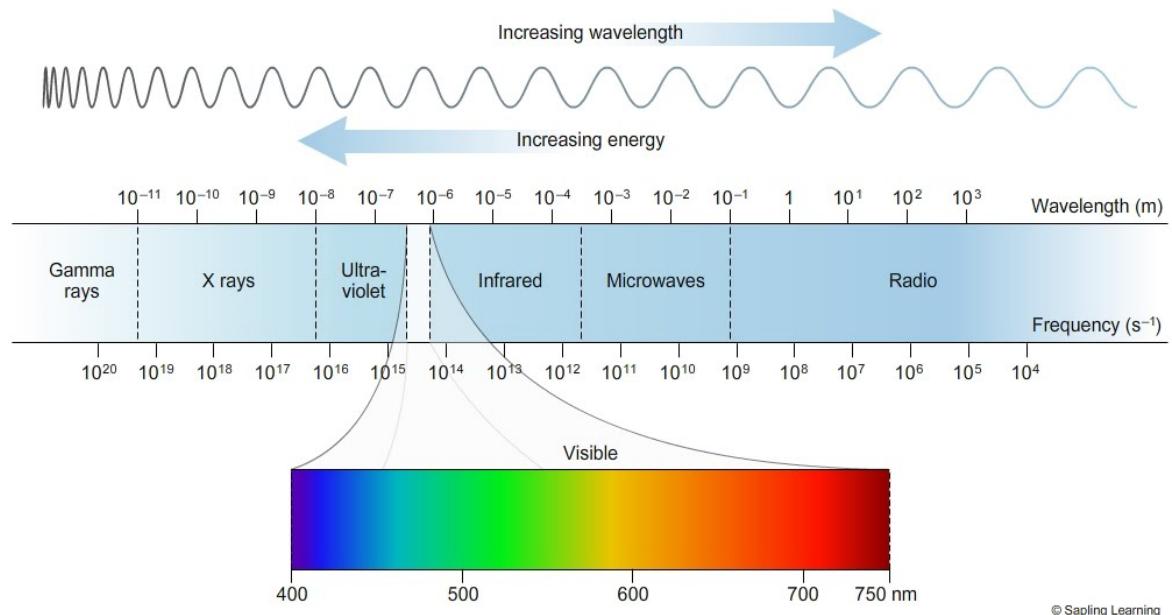


Fig. 1.1 Light spectrum [1]

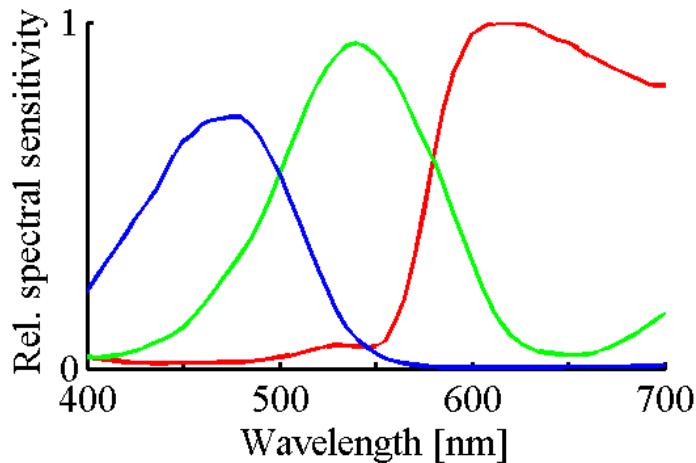


Fig. 1.2 Spectral response of an RGB camera[2]

human eye and common RGB camera filters, many details are lost in narrow parts of the electromagnetic spectrum. These details however, can be quite meaningful. Subtle difference in spectra can indicate important details in material properties. This is a property that exploited both in nature and scientific applications. The term “Multispectral Imaging” broadly refers to any sensing systems that utilize more than the standard RGB bands of sensitivity. Therefore, a Multispectral Images consists of several monochrome images which are corresponding specific wavelength or a wavelength band from the electromagnetic spectral and expands the coverage beyond the visible range. Since, a Multispectral image contains spectral information beyond the visible range, it can provide more information than an RGB image.

1.2 Applications of Multispectral Imaging

Multispectral imaging has largely been used for remote sensing, surveillance, and industrial applications, where the spectral signatures are used to differentiate between materials. Examples of industry applications include[7],

- **Health-care:** The earliest, most successful uses of MSI were initially in diagnostic medicine. Multispectral imaging allows health-care providers to pinpoint the presence of diseases hard to identify through other means. Soon, MSI may be combined with nano-technology to diagnose health issues at the level of individual cells.
- **Food industry:** The multispectral imaging technique is currently taking many challenges to be accepted as the most preferable analytical tool in identifying compositional

fingerprints of food products and their authentication. The need for fast and reliable method of authenticity and object identification has increased the interest in the application of multispectral imaging for quality control in the agricultural, pharmaceutical, and food industries.

1.3 Terminology

1.3.1 Spectral Range

The spectral range describes the wavelength regions covered by the multispectral imaging system. Spectral imaging instruments could cover either the ultraviolet (UV), visible, near-infrared or infrared wavelengths based on the required application. Multispectral imaging system in the visible and very near-infrared range 380nm-800nm or 400nm-1000nm is the most widely used.

1.3.2 Spectral Resolution

The spectral resolution is defined as the absolute limit of the ability of a multispectral imaging system to separate two adjacent monochromatic spectral features emitted by a point in the image. Spectral resolution is the measure of the narrowest spectral feature that can be resolved by a multispectral imaging system.

1.3.3 Spatial Resolution

The spatial resolution of the multispectral imaging system determines the size of the smallest object that can be seen on the surface of the specimen by the sensor as a distinct object separate from its surroundings. Spatial resolution also determines the ability of a system to record details of the objects under study. Higher spatial resolution means more image detail expanded.

1.3.4 Band Numbers

The number of bands is one of the main parameters that characterize multispectral imaging systems. Based on the type of spectral imaging system (i.e. Multispectral and Hyperspectral), the number of spectral bands could vary from a few (usually fewer than 10) in multispectral imaging to about 100-250 spectral bands in the electromagnetic spectrum in the case of hyperspectral imaging.

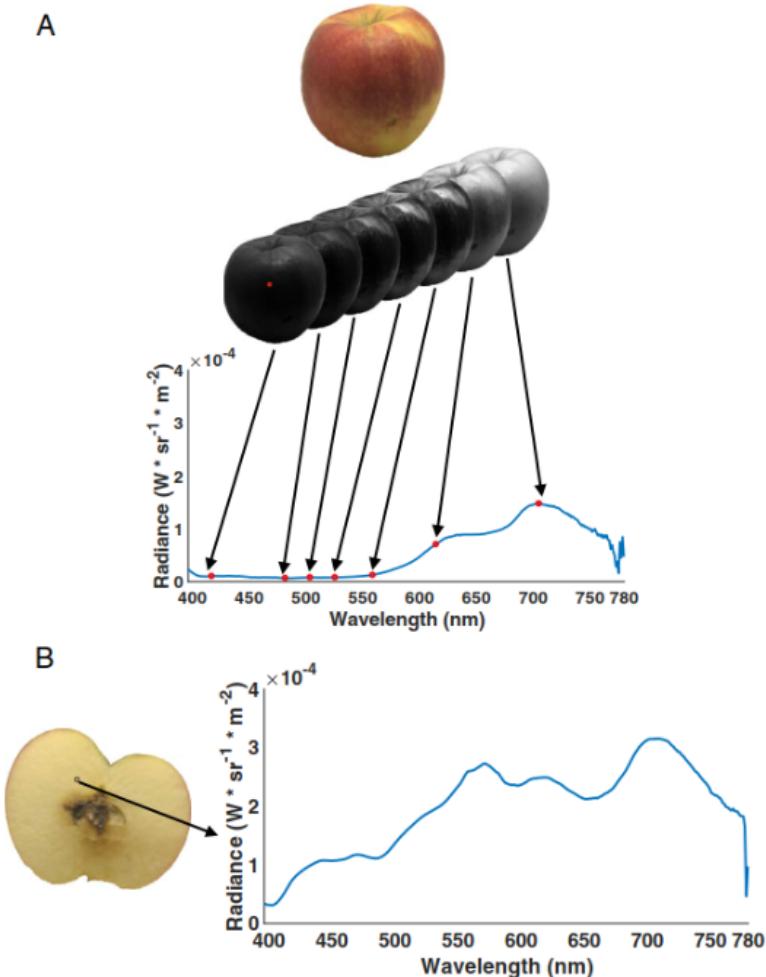


Fig. 1.3 Idea of spectral signature[3]

1.3.5 Spectral Signature

Multispectral imaging exploits the fact that all materials, due to the difference of their chemical composition and inherent physical structure, reflect, scatter, absorb, and/or emit electromagnetic energy in distinctive patterns at specific wavelengths. This characteristic is called spectral signature. Every pixel in the multispectral image contains its own spectral signature. Briefly, spectral signature is defined as the pattern of reflection, absorption, transmittance, and/or emitting the electromagnetic energy at specific wavelengths. In principle, Spectral signature can be used to uniquely characterize, identify, and discriminate by class/type and given object(s) in an image over a sufficient broad wavelength band (see Fig. 1.3).

1.4 Representation of Multispectral Image Data

1.4.1 The Data-cube

Multispectral data consists of several monochrome images representing intensities at different wavelength bands composed of vector pixels (voxels) containing two-dimensional spatial information (of m rows and n columns) as well as spectral information (of K wavelengths). These data are known as a three-dimensional multispectral cube, or data-cube (see Fig. 1.4), which can provide physical and/or chemical information of a material under test. This information can include physical and geometric observations of size, orientation, shape, color, and texture, as well as chemical information[8]. However, the combination of these two features (spectral and spatial) is not trivial, mainly because it requires creating a three-dimensional (3D) data set that contains many images of the same object, where each one of them is measured at a different wavelength. Because pixels are digitalized gray values or intensities at a certain wavelength, they may be expressed as integers. Intensity values of a spatial image in the data-cube at one wavelength may have 8-bit gray values meaning that 0 is the black and 255 is the white.

Figure 1.4 shows one example of the data-cube extracted from a multispectral image acquired for a piece of meat. The raw multispectral image consists of a series of contiguous sub-images; each one represents the intensity and spatial distribution of the tested object at a certain waveband. All individual spatial images could be picked up from the data-cube at any wavelength(s) covering the spectral sensitivity of the system. Therefore, a multispectral image described as $I(x, y, \lambda)$ can be viewed either as a separate spatial image $I(x, y)$ at each wavelength (λ), or as a spectrum $I(\lambda)$ at every pixel (x, y) . Each pixel in a multispectral image contains the spectrum of that specific position. The resulting spectrum acts like a signature which can be used to characterize the composition of that particular pixel. Since multispectral imaging acquires spatially distributed spectral response at pixel levels, this allows flexible selection of any region of interest on a target object. For instance, if two different pixels from two different compositional locations in a data-cube are extracted, they will show different spectral signatures. Therefore, without any further manipulation or preprocessing treatments of these spectral data, the difference in spectral signatures between lean meat pixel and fat met pixel of the tested piece of meat shown in Fig. 1.4 are noticeably distinguished.

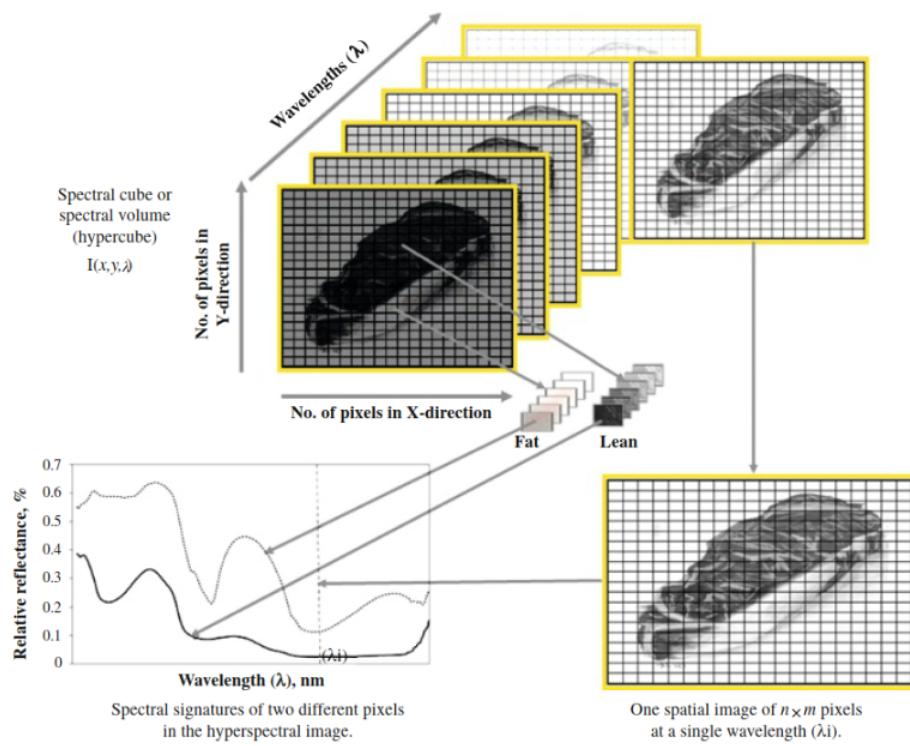


Fig. 1.4 The schematic diagram of multispectral image (data-cube) for a piece of meat showing the relationship between spectral and spatial dimensions. Every pixel in the multispectral image is represented by an individual spectrum containing information about chemical composition at this pixel[4].

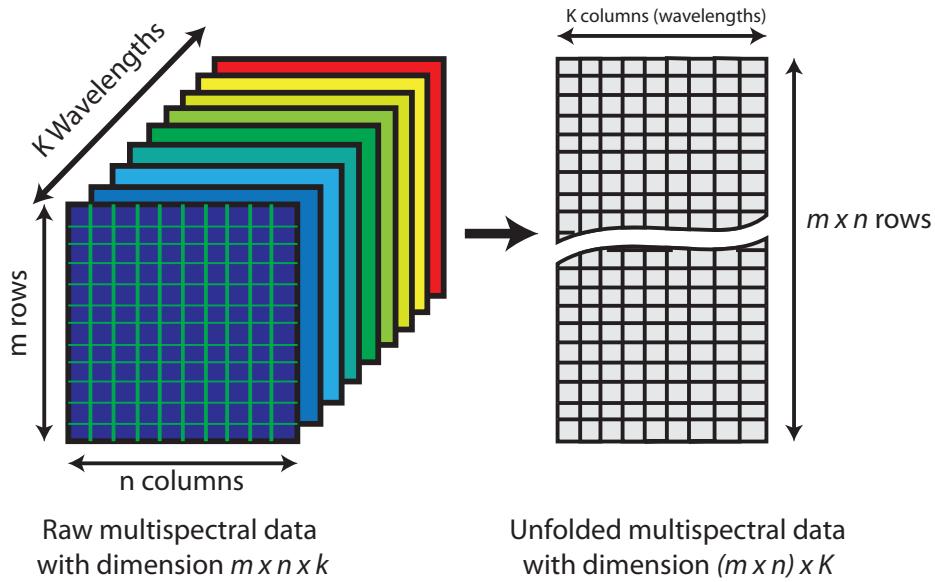


Fig. 1.5 Unfolding multispectral data "data-cube"

1.4.2 Features of Multispectral Data - "data-cube"

- **Multispectral data are inherently high dimensional** since they are, by definition, composed of several number of spectral bands (usually less than 10).
- **The data-cube** can be viewed in the spatial domain as images ($m \times n$) at different wavelengths or in the spectral domain as spectral vectors at all wavelengths, as shown in Fig 1.5. Both representations are useful when analyzing multispectral data using statistical tools. For an example, if one multispectral image has dimension of $480 \times 640 \times 6$, this data-cube can be interpreted as six single channel images each with 480×640 pixels. Alternatively, the same data-cube can be viewed as $480 \times 640 = 307200$ spectra, each with six wavelength points. This huge amount of data poses data mining challenges, but also creates new opportunities for discovering detailed hidden information.

Chapter 2

Statistical Tools Used for Analysis of Multispectral Data

2.1 Introduction

Because multispectral imaging technology provides spectral information of several spectral bands (usually less than 10 spectral bands), an effective approach for data analysis and pattern classification is necessary to extract the desired information from images. Much work has been carried out in the literature to present the feature extraction and pattern recognition methods in multispectral image classification.

A general two step strategy, which is feature extraction followed by pattern classification. The feature extraction step is also called optimum bands selection or extraction, whose aim is to reduce or transform feature space into another space of a lower dimensionality. **Principal Component Analysis (PCA)** , **Fisher's Discriminant Analysis (FDA)** and/or **Feature Space Discriminant Analysis (FSDA)** followed by **K-means** clustering is the most popular technique in this method.

2.2 Statistical Tools used for Feature Extraction and Dimensionality Reduction

In multispectral image analysis the data dimension is about six to ten. Hence, it is necessary to reduce the data redundancy and efficiently represent the distribution of data. In order for reduce the dimension and to extract the features following statistical tools were used.

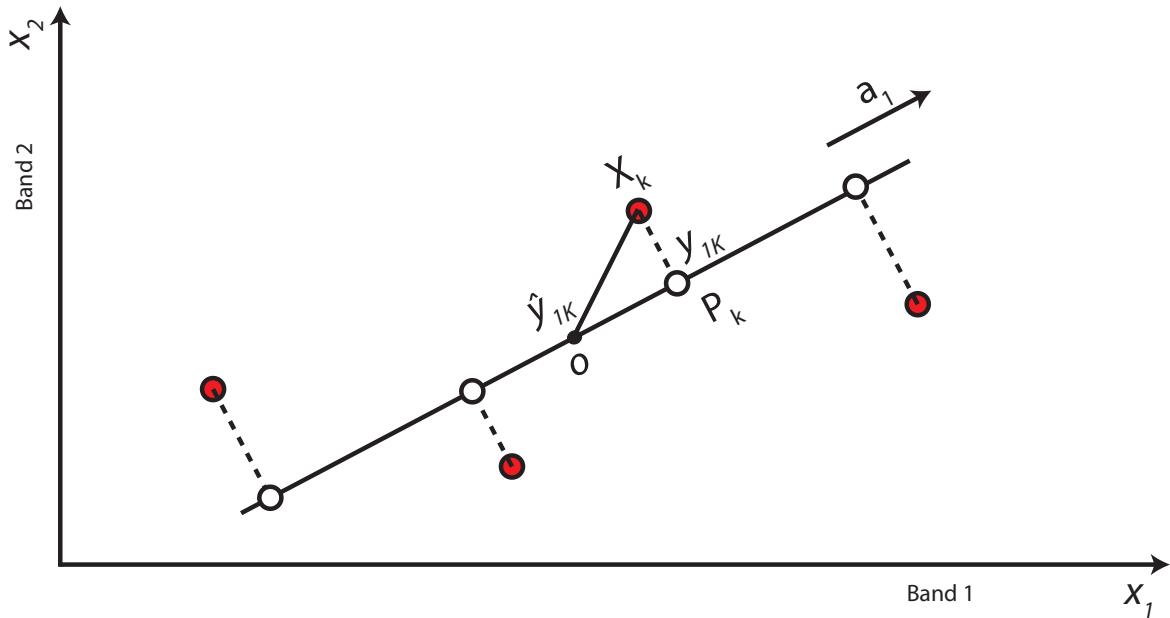


Fig. 2.1 Geometrical illustration of PCA as an orthogonal projection to a lower-dimensional subspace. Points in a plane are projected onto a line.

2.2.1 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a special linear transformation that is widely used for applications like dimensionality reduction, data compression, feature extraction and data visualization. Principal components were first proposed by person in 1901 and further developed by Hotelling in 1933.

The Basic Idea of PCA

The basic idea of PCA[9–13] is illustrated in Fig. 2.1 with two-dimensional data. Now we need to determine what is the best line through the set of data points denoted by the red circles? To answer this question we must clarify what is meant by "best"? A line is a one-dimensional subspace specified by a unit vector a_1 . The orthogonal projection of the observation x_k onto the line is point $y_{1k} = a_1^T x_k$. PCA determines the best line by minimizing the sum of the squares of the perpendicular distances $\sum_k (A_k P_k)^2$ of the data points from the line. Note that $A_k P_k$ is the error we make when we approximate the data point A_k by its projection P_k . The mean value of the projections y_{1k} is $\bar{y}_1 = (1/N) \sum_k a_1^T x_k = a_1^T \hat{m}$: that is, the projection of the data mean onto the line. Thus, the variance of the projection y_{1k} is proportional to $\sum_k (OP_k)^2$. From the orthogonal triangle $OP_k A_k$ and the Pythagorean theorem,

we have

$$\sum_k (OA_k)^2 = \sum_k (OP_k)^2 + \sum_k (A_k P_k)^2. \quad (2.1)$$

Since the sum $\sum_k (OA_k)^2$ is fixed for a given swarm of points, minimization of mean square error $(1/N) \sum_k (A_k P_k)^2$ is equivalent to minimization of variance $(1/N) \sum_k (OP_k)^2$ under the constraint $a_1^T a_A = 1$. The solution to this problem is described in the next section.

The Steps of PCA

Step 1: Calculate the mean vector (\hat{m}) from the available data $x_1, x_2, x_3, \dots, x_N$,

$$\hat{m} = \frac{1}{N} \sum_{k=1}^N x_k \quad (2.2)$$

where:

N = Number of available data points

x_k = k^{th} data vector

Step 2: Calculate the covariance matrix (C) given by,

$$C = \frac{1}{N} \sum_{k=1}^N (x_k - \hat{m})(x_k - \hat{m})^T \quad (2.3)$$

where:

N = Number of available data points

x_k = k^{th} data vector

\hat{m} = Mean vector

Step 3: Calculate the eigenvectors v_k and eigenvalues λ_k of C ,

$$|C - \lambda I| = 0; \text{ to find eigenvalues of } C \quad (2.4)$$

where:

C = Covariance matrix

λ = Eigenvalue of the covariance matrix

I = Identity matrix

$$Cv = \lambda v; \text{ to find eigenvectors of } C \quad (2.5)$$

where:

C = Covariance matrix

λ = Eigenvalue of the covariance matrix

v = Eigenvector of C

Step 4: Deriving the new data set y ,

$$y = [v_1 \ v_2 \ v_3 \ ... \ v_p]^T x \quad (2.6)$$

where:

y = Transformed data

$v_i = i = 1, 2, 3, \dots, p$; Eigenvectors of covariance matrix (C)

I = Identity matrix

2.2.2 Fisher's Discriminant Analysis (FDA)

PCA makes **no** a priori assumptions about the existence of classes in the data; even if class information is available, it is not used. Fisher's discriminant analysis on the other hand, assumes that every N observation vectors x_1, x_2, \dots, x_N are divided a priori into K classes with N_k observations in k^{th} class ζ_k , where $N = N_1 + N_2 + \dots + N_K$. In the presence of class structure, the best type of linear transformation will be dictated by the objective of the analysis. In general, we use PCA to highlight *similarities* in the data and linear discriminant analysis to highlight *differences* between the classes.

By using this algorithm, the original dataset can be transformed to a new space where the variation among the classes is maximized while the variation within a class is minimized. In other words, there is a considerable difference among the classes and the differences within a class are smaller in the new space, thus enabling easy identification of different clusters. Algorithm for FDA is described below.

Steps of FDA Algorithm

Step 1: Construction of between class scatter matrix and within class scatter matrix. The equations to construct between class scatter matrix and within class scatter matrix

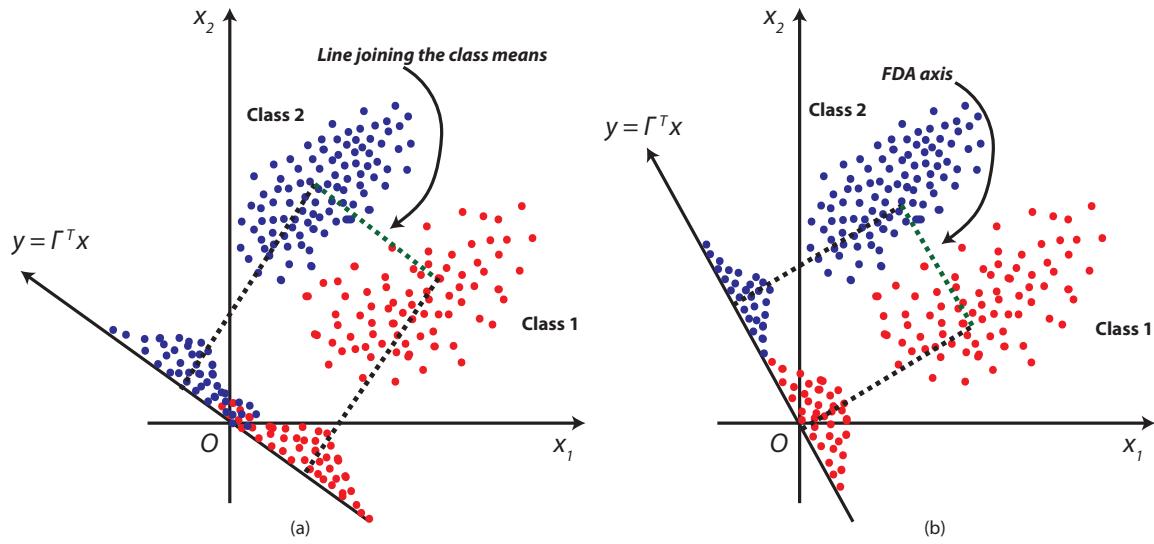


Fig. 2.2 (a): Samples from two classes projection onto the line joining the class means. Note that there is considerable class overlap in the projected space. (b): Corresponding projection based on the Fisher linear discriminant (FDA), showing the greatly improved class separation.

are given below. Our aim is to transform the dataset in to a new space where between class scatter is maximized and within class scatter is minimized.

$$\begin{aligned}
 S_b &= \sum_{i=1}^{i=l} n_i (\mu_i - \mu)(\mu_i - \mu)^T \\
 S_w &= \sum_{i=1}^l \sum_{j \in i} (x_j - \mu_i)(x_j - \mu_i)^T \\
 \mu_i &= \frac{a}{n_i} \sum_{j \in i} x_j \\
 \mu &= \frac{1}{n} \sum_{i=1}^n x_i
 \end{aligned} \tag{2.7}$$

where:

S_b = between class scatter matrix,
 μ = vector representing mean spectral information of all pixels,
 μ_i = vector representing mean spectral information of the pixels of class i,
 n_i = number of pixels in the training sample belonging to class i,
 l = number of classes,
 S_w = within class scatter matrix,
 x_j = vector representing the spectral information of a pixel in the training sample belonging to class j

Step 2: Constructing the Transformation Matrix

$$S_b \varphi = \lambda S_w \varphi \quad (2.8)$$

where:

S_b = between class scatter matrix
 S_w = within class scatter matrix
 λ = eigenvalues of the fisher matrix, $S_w^{-1} S_b$
 φ = eigenvectors of the fisher matrix, $S_w^{-1} S_b$.

The transformation matrix that transforms the original data to a new space with reduced dimensionality is given by,

$$\Gamma = [\varphi_1 \varphi_2 \varphi_3 \dots \varphi_n]^T \quad (2.9)$$

where:

Γ = transformation matrix
 $\varphi_1 \varphi_2 \varphi_3 \dots \varphi_n$ = real eigenvectors of fishers matrix

Step 3: Transforming the New Space

The dataset is transformed to the new space using,

$$V_2 = \Gamma V_1 \quad (2.10)$$

where,

V_1 = matrix representing all the pixels in the original space,
 V_2 = matrix representing the pixels in the new space.

2.2.3 Feature Space Discriminant Analysis (FSDA)

Popular feature extraction methods such as Linear Discriminant Analysis (LDA), Generalized Discriminant Analysis (GDA), Nonparametric Weighted Feature Extraction (NWFE), and Median Mean-line Discriminant Analysis (MMLDA) just use the class discrimination for feature extraction. In FSDA, the difference between spectral bands in the transformed feature space, in addition to separability between classes is considered. In other words, the features in FSDA are extracted in such a way that[14]:

- The produced features (spectral bands) are as different from each other as possible
- Separability between classes is increased

To fulfill the above two tasks, a training set of sample set is used, in which the class of each sample point is known. Each spectral band is considered as a point in the Euclidean space where the axes represent the classes. This is quite unusual as the normal practice is to represent each pixel as a point in the Euclidean space where the axes represent spectral bands. This representation is used in order to transform the original spectral bands into a new set that maximizes the difference between the spectral bands. Hence, the first step of FSDA is achieved by the equations,

$$\begin{matrix} M_1 & M_2 & M_3 & \dots & M_c \\ \left(\begin{array}{ccccc} m_{11} & m_{12} & m_{13} & \dots & m_{1c} \\ m_{21} & m_{22} & m_{23} & \dots & m_{2c} \\ m_{31} & m_{32} & m_{33} & \dots & m_{3c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{d1} & m_{d2} & m_{d3} & \dots & m_{dc} \end{array} \right) h_1 \\ h_2 \\ h_3 \\ \vdots \\ h_d \end{matrix} \quad (2.11)$$

where, m_{ij} is the mean of class j and spectral feature i , c and d are the number of classes and the number of spectral bands (features) respectively. The entries of the columns and rows of matrix in 2.11 are,

$$\begin{aligned} h_i &= [m_{i1} \ m_{i2} \ m_{i3} \ \dots \ m_{ic}]^T, \quad i = 1, 2, 3, \dots, d \\ M_j &= [m_{1j} \ m_{2j} \ m_{3j} \ \dots \ m_{dj}]^T, \quad j = 1, 2, 3, \dots, c \end{aligned} \quad (2.12)$$

where, h_i is the mean of c classes in i^{th} dimension and M_j is the mean of class j in d dimensions. In other words, h_i is the representative of band (feature) i and M_j is the representative of class j . The first projection matrix in FSDA method should be found such that, the between-spectral scatter (to produce features as different from each other as possible)

is maximized. This requires the calculation of the between-spectral scatter matrix, S_f given by,

$$S_f = \sum_{i=1}^d d(h_i - \bar{h})(h_i - \bar{h})^T \quad (2.13)$$

Once the transformation matrix is obtained from the eigenvectors of the matrix S_f , the h_i vectors (vector representing each spectral band) are multiplied by the transformation matrix, to obtain the spectral bands in the new space that maximizes the difference between the spectral bands. The transforming equation and the matrix containing the new spectral bands are given by,

$$(g_i)_{c \times 1} = W_{c \times c}(h_i)_{c \times 1}, \quad i = 1, 2, 3, \dots, d \quad (2.14)$$

$$\begin{array}{ccccc} R_1 & R_2 & R_3 & \dots & R_c \\ \left(\begin{array}{ccccc} r_{11} & r_{12} & r_{13} & \dots & r_{1c} \\ r_{21} & r_{22} & r_{23} & \dots & r_{2c} \\ r_{31} & r_{32} & r_{33} & \dots & m_{3c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{d1} & m_{d2} & m_{d3} & \dots & m_{dc} \end{array} \right) & g_1 & & & \\ & & g_2 & & \\ & & g_3 & & \\ & & & & \vdots \\ & & & & g_d \end{array} \quad (2.15)$$

where, $W_{c \times c}$ is the transformation matrix obtained by the dominant eigenvectors of S_f , g_i 's in the new space are the corresponding vectors of h_i 's in the original space and R_i 's in the new space are the corresponding vectors of M_i 's in the original space.

In the new feature space, the difference between spectral bands is increased. In other words, g_i ; ($i = 1, 2, \dots, d$) are more different from each other than h_i ; ($i = 1, 2, \dots, d$). In the second step of FSDA, the between class scatter is maximized increasing the class discrimination. The between-class scatter matrix (S_b) is calculated using,

$$S_b = \sum_{j=1}^c (R_j - \bar{R})(R_j - \bar{R})^T \quad (2.16)$$

where,

$$\bar{R} = \frac{1}{c} \sum_{j=1}^c R_j \quad (2.17)$$

For the extraction of p number of features from the original $d \times 1$ feature vector, x , p number of eigenvectors of S_b associated with p largest eigenvalues of it compose the projection matrix, A and the transformation is given by,

$$y_{p \times 1} = A_{p \times p} x_{d \times 1} \quad (2.18)$$

where, x is a pixel in the original d dimensional space (with reflectance values of d spectral bands) and y is the same pixel represented in a new p dimensional space with spectral information of p number of spectral bands, where, the spectral scatter and class scatter is maximized for the convenience of classification.

2.3 Tools used for Clustering

2.3.1 Spectral Clustering

Introduction

Clustering is one of the most widely used techniques for exploratory data analysis in statistics. Clustering is important when identifying similar behavior in a data set. It is proven that the results obtained by spectral clustering very often outperform the traditional approaches. Spectral clustering is very simple to implement and can be solved efficiently by standard linear algebra methods. And also it is a method of clustering data using the "Spectrum" (Eigenvalues) of the affinity matrix. There are many spectral clustering algorithms available, but in the project normalized spectral clustering algorithm is implemented. The spectral clustering algorithm is as follows[12, 15].

Algorithm Steps

Step 1: Constructing Disparity Matrix Disparity matrix is defined as,

$$D_{i,j} = ||S_i - S_j|| \quad (2.19)$$

where, the $(i, j)^{th}$ element of the disparity matrix denotes the Euclidian distance between case S_i and S_j . In this step what is done is taking the Euclidian distance of all vectors with all the other vectors. This gives a measurement of comparison among all the vectors.

Step 2: Finding the affinity matrix "Affinity" defines how close or similar are two points given in the feature space. Thus trajectories belonging to the same class should have a high affinity whereas different classes should have a low affinity. This is achieved by transforming the Disparity matrix through Gaussian Kernel. Due to the exponential curve differences between small disparity values are amplified. Hence, clustering is more accurate. Here determining the best sigma is an important issue.

Analysis of the significance of sigma is discussed in a latter chapter. The affinity matrix is formed using,

$$A(i, j) = e^{-\frac{D(i, j)^2}{2\sigma^2}} \quad (2.20)$$

where,

A = Affinity matrix

D = Disparity matrix

σ = Tunable constant

Step 3: Construct the diagonal matrix, W The $(i, j)^{th}$ element of this diagonal matrix is the sum of all the elements in i^{th} row of affinity matrix given by,

$$W(i, j) = \sum_{\forall j} A(i, j) \quad (2.21)$$

Step 4: Construct the Laplacian Matrix, L given by,

$$L = W^{-\frac{1}{2}} A W^{\frac{1}{2}} \quad (2.22)$$

Step 5: Constructing the matrix, X The first prominent k eigenvectors ($x_i, i = 1, 2, 3, \dots, k$) of the Laplacian matrix was taken as X ,

$$X = [x_1 \ x_2 \ x_3 \ \dots \ x_k] \quad (2.23)$$

The rows of the matrix, X gives the pixel coordinates of the new space which have been clustered.

2.4 Tools used for Classification

2.4.1 K-Means

Introduction

K-means clustering is a type of unsupervised learning, which is used when there is unlabeled data (i.e., data without defined categories or groups). The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable K . The algorithm

works iteratively to assign each data point to one of K groups based on the features that are provided. Data points are clustered based on feature similarity. The results of the K-means clustering algorithm are:

- The centroids of the K clusters, which can be used to label new data
- Labels for the training data (each data point is assigned to a single cluster)

The procedure is about classifying a given data set using a certain number of clusters. (k clusters) To do that k centers should be defined, one for each cluster. These centers should be placed as far as possible from each other. Next, each point belonging to the data set is allocated to the nearest center. After that it is required to calculate new centroids to each cluster. After taking these k new centroids, a new binding has to be done between the same data set points and the nearest new center. A loop has been generated. As a result of this loop it could be noticed that the k centers change their location step by step until no more changes are done or in other words centers do not move any more. Finally, this algorithm aims at minimizing an objective function, known as squared error function given by,

$$J = \sum_{i=1}^c c_i \sum_{j=1}^{c_i} ||x_j - v_i||^2 \quad (2.24)$$

where,

$||x_i - v_i||$ = the Euclidian distance between x_i and v_i

c_i = the number of data points in i^{th} cluster

c = the number of cluster centers.

Algorithm steps for K-means clustering Let,

$$X = [x_1, x_2, x_3, \dots, x_n] \quad (2.25)$$

be the first data points and

$$V = [v_1, v_2, \dots, v_c] \quad (2.26)$$

be the set of centers.

Then,

- Randomly select ' c ' cluster centers.
- Calculate the distance between each data point and cluster centers.

- Assign the data point to the cluster center whose distance from the cluster center is the minimum of all the cluster centers. Recalculate the new cluster center using,

$$v_i = \frac{1}{c_i} \sum_{j=1}^{c_i} x_i \quad (2.27)$$

where, ' c_i ' represents the number of data points in the i^{th} cluster.

- Recalculate the distance between each data point and new obtained cluster centers.
- If no data point was reassigned then stop, otherwise repeat from step 3.

Chapter 3

Analysis of Multispectral Images Taken from the Database

This is the place where we have begun our analysis. We were developed MATLAB codes for existing algorithms (described in the previous chapter) and applied them to multispectral images taken from the Internet. At the same time, we have developed our own multispectral imaging system.

3.1 About the Database

The database used for multispectral analysis is one of the projects conducted in Columbia University, Department of Computer Science[5]. Database consists of multispectral images that were used to emulate the GAP camera[16]. The images are of a wide variety of real-world materials and objects.

These multispectral images represent the reflectance of the materials in the scene. The reflectance spectrum at each pixel was computed from the measured multispectral image using calibrated estimates of the illumination spectrum and camera spectral response. As a result, each multispectral image should be taken as a close approximation, rather than exact measurement, of the true reflectance of the scene.

The database consists of 32 scenes, divided into 5 sections. Each scene has an associated zip file. These zip files include full spectral resolution reflectance data from 400nm to 700nm at 10nm steps (31 bands total). Each band is stored as a 16-bit gray scale PNG image. Each scene also contains a single representative color image, displayed using RGB values rendered under a neutral daylight illuminate(D65). A multispectral image of a fake and real lemon slices is shown in Fig. 3.1.

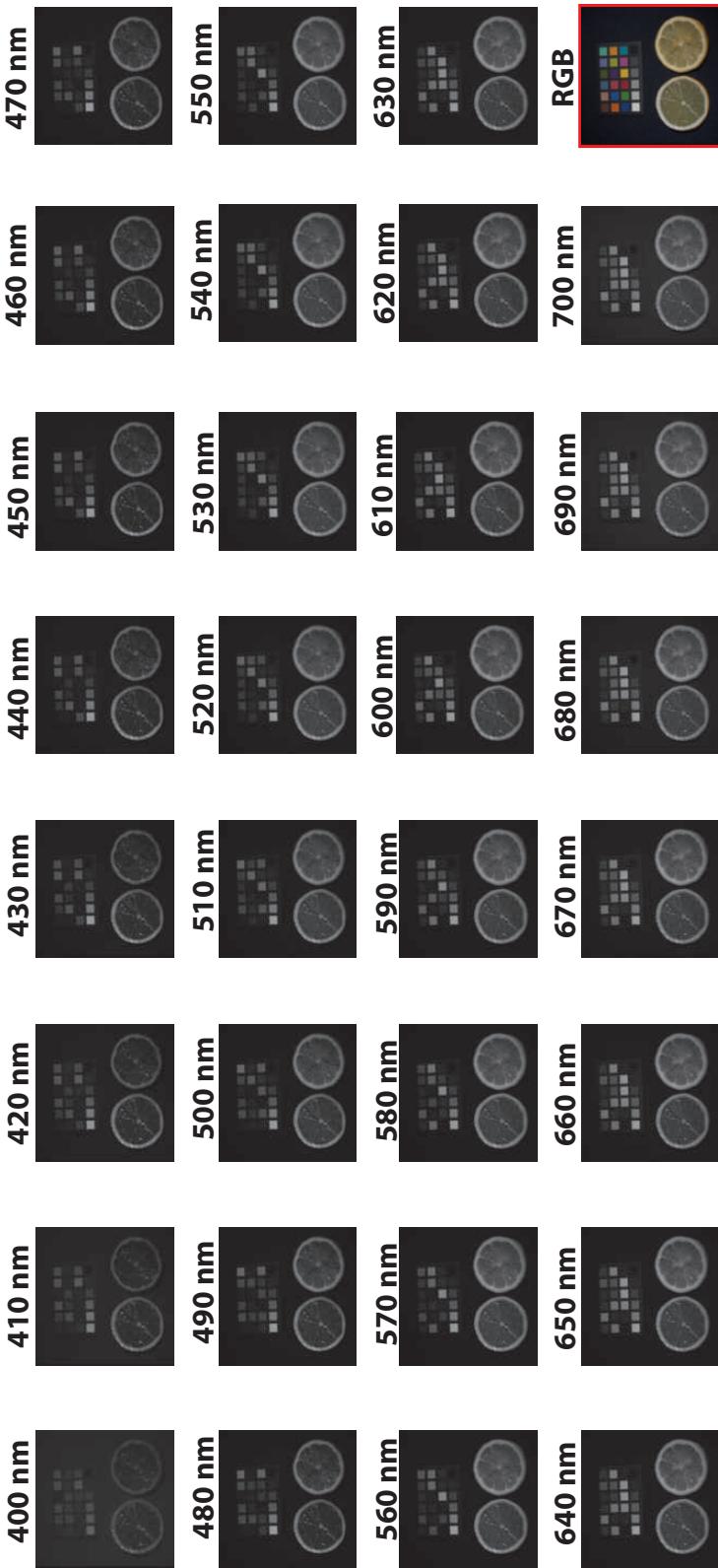


Fig. 3.1 Multispectral image of a fake and real lemon slices taken from the database[5] for our analysis

Table 3.1 Details of multispectral images database[5]

Camera	Cooled CCD camera (Apogee Alta U260)
Resolution	512 x 512 pixel
Filter	VariSpec liquid crystal tunable filter
Illuminant	CIE Standard Illuminant D65
Range of wavelength	400nm - 700nm
Steps	10nm
Number of band	31 band
Focal length	$f/1.4$
Focus	Fixed (focused using 550nm image)
Image format	PNG (16bit)

3.2 Analysis of Multispectral Images

Multispectral images of fake and real strawberries were chosen for this analysis as shown in Fig.??.

3.2.1 Selection of pixels from fake and real strawberries

121 pixels from each fake and real strawberries were selected for this analysis as shown in Fig. 3.2.

3.2.2 Spectral Signatures of all the Pixels

After the pixels were selected from fake and real strawberries, the intensity values of each pixel were plotted with wavelength to obtain spectral signatures as shown in Fig. 3.3. It is clearly see that there is no much deference in spectral signatures of fake and real strawberries from the spectral signature plot. Therefore, in order to differentiate fake and real strawberries, we were applied the Principal Component Algorithm (PCA) to the dataset.

3.2.3 Applying Principal Component Analysis (PCA)

Our next step was to perform Principal Component Analysis (PCA) on the dataset of fake and real strawberries. The algorithm of PCA was described in the previous chapter. Then we were plotted the new spectral signatures of fake and real strawberries for first ten principal components as shown in Fig. 3.4. It is clear that the variation of spectral signatures are more



Fig. 3.2 Selected pixels from fake and real strawberries (121 pixels from fake strawberries and 121 pixels from real strawberries.)

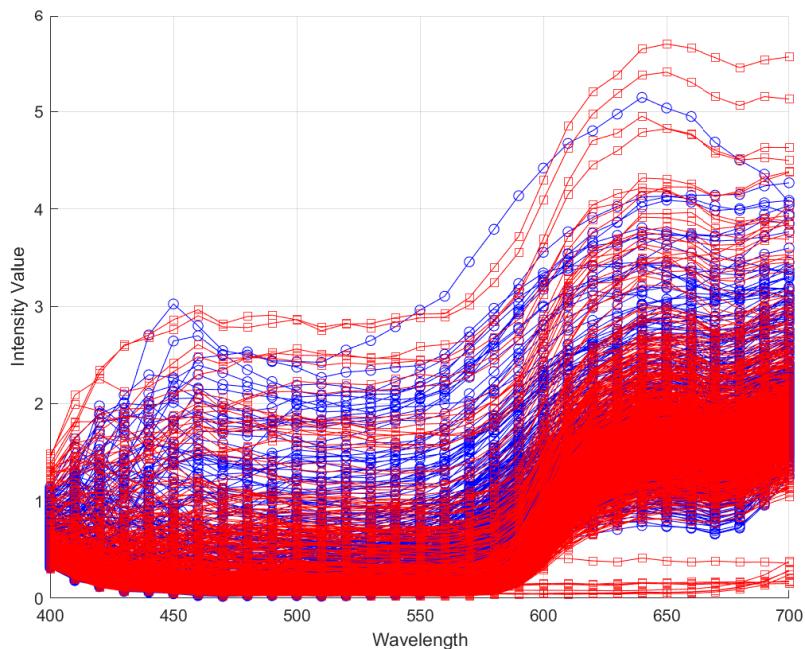


Fig. 3.3 Spectral signatures of all the selected pixels from fake and real strawberries - *Red* lines for fake strawberries and *blue* lines for real strawberries. The x-axis is wavelength(nm) and y axis is intensity value

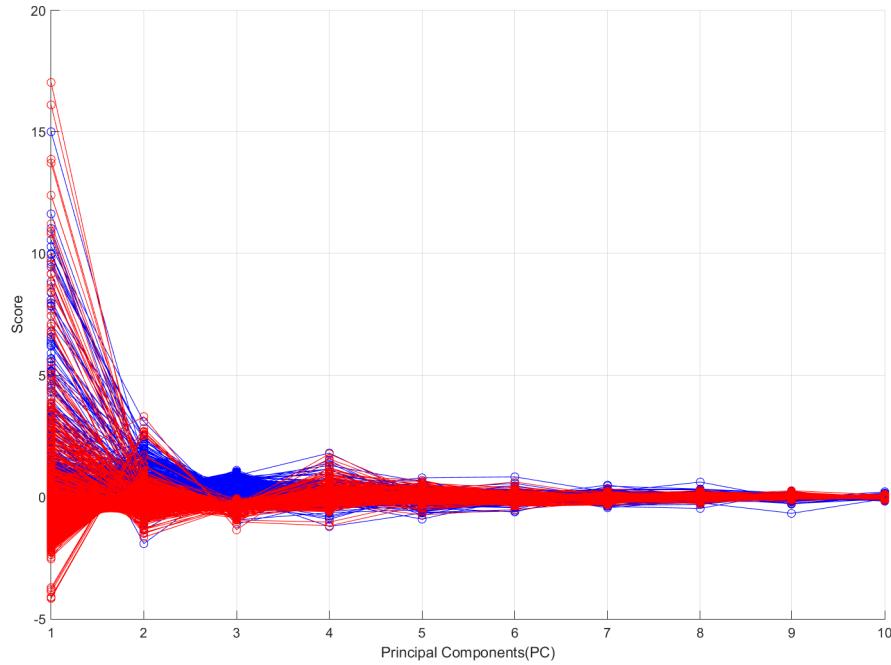


Fig. 3.4 Spectral signatures of fake and real strawberries in reduced dimensional space (for first ten principal components) - *Red* lines for fake strawberries and *blue* lines for real strawberries.

dominant in first few principal components whereas variation of spectral signatures are more week for last principal components.

3.2.4 Scatter Plot of All the Pixels in Reduced Dimensional Space

Once the Principal Component Analysis was performed on the dataset of fake and real strawberries then we were plotted the scatter plot against first three principal comportsments as shown in Fig 3.5. From that graph we can clearly differentiated fake strawberries from real strawberries. But in order to apply a classification algorithm such as K-means, we should need to increase the distance between inter classes while reducing the intra-class distance. For that we were performed the spectral clustering algorithm on the PCA data.

3.2.5 Applying Spectral Clustering Algorithm

In order to improve the classification accuracy of K-means algorithm, we were performed the spectral clustering algorithm on the dataset taken from after performing the principal component analysis. As the spectral clustering algorithm described in the previous chapter, we were plotted the eigenvalue difference between third and second, fourth and third, fifth

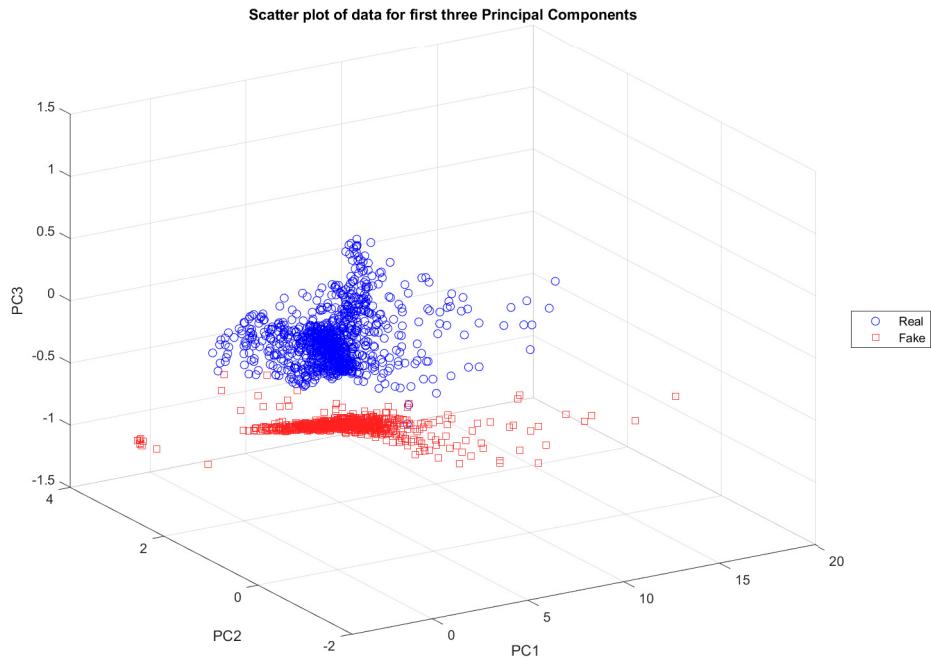


Fig. 3.5 Scatter plot of fake and real strawberries in reduced dimensional space (PC1, PC2 and PC3) - *Red* circles for fake strawberries and *blue* circles for real strawberries.

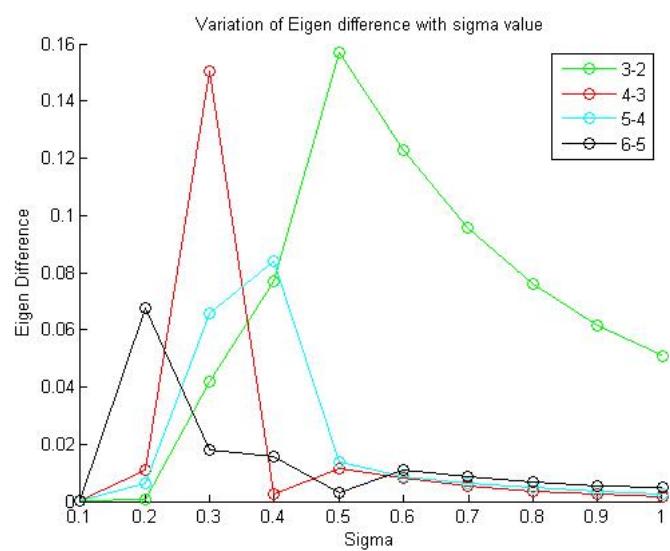


Fig. 3.6 Variation of eigenvalue difference for different sigma(σ) values

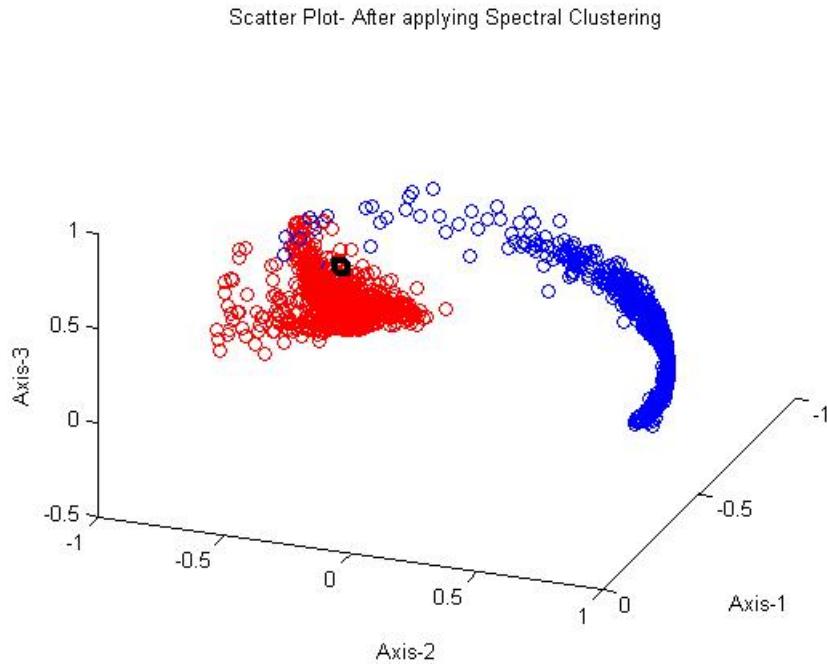


Fig. 3.7 Scatter plot of fake and real strawberries after applying spectral clustering algorithm to data from PCA - *Red* circles for fake strawberries and *blue* circles for real strawberries.

and fourth, and sixth and fifth for different sigma(σ) as shown in Fig. 3.6. Since the highest variation of eigenvalue was obtained for eigenvalue difference between third and second eigenvalues at $\sigma = 0.5$, that value of sigma was used as the tunable parameter in the spectral clustering algorithm. Finally, the scatter plot of fake and real strawberries after applying the spectral clustering algorithm was shown in Fig. 3.7.

3.2.6 Classification of real and fake lemon slices using K-means Algorithm

Finally, the dataset of fake and real strawberries were classified with the aid of K-means algorithm. The test results were shown in Tab. 3.2. Therefore, class 1 can be identified as real strawberries and class 2 can be identified as fake strawberries.

Table 3.2 Results of classifying fake and real strawberries using k-means algorithm after performing PCA and spectral clustering

	class 1	class 2
Real	116	5
Fake	0	121

Chapter 4

Hardware Design of the Multispectral Imaging System

Multispectral cameras with high spectral and spatial resolutions are available in the market but are quite expensive. Therefore, implementing such a multispectral imaging system for the use of a small or a medium scale industry is not commercially viable. At the same time, depending on the application it may not require a multispectral camera with higher spectral and spatial resolution that eventually lead to a higher cost and computational complexity. Hence a low-cost multispectral imaging system was developed using a smartphone RGB camera[17, 18]. This imaging system has the capability of capturing multispectral images from Ultra-Violet (UV) to Near Infra-Red (NIR) having a resolution of six spectral bands¹ (see Tab. 6.1 and Fig. 4.1).

4.1 Selection of Spectral Bands

For this implementation, a total of ten narrow band LEDs with wavelengths ranging from 405 nm to 960 nm were selected. Details of the narrow-band LEDs are given in Tab. 6.1.

These were off-the-shelf LEDs, and the intensities of all the LEDs were adjusted to a constant approximated value (as the power spectral distribution of each LED in Fig. 4.1) using a programmable LED driver IC (MAX16839ASA+). Then all the LEDs were placed on

¹Actually, this multispectral imaging system comprised of ten narrow-band LEDs (as given in the Tab 6.1) which are ranging from 405 nm to 960 nm. Due to the presence of IR filter implemented in-front of the smartphone RGB sensor, we are able to see images up to 740 nm with resolution of six spectral bands. In other words images correspond to 850 nm, 890 nm, 950 nm and 960 nm are not clearly visible from the smartphone camera. So, the effective range of this multispectral imaging system with smartphone camera (as the sensing device) is from 405 nm to 740 nm with resolution of six spectral bands even though the implemented system contains ten narrow-band LEDs.

Table 4.1 Details of Narrow-band LEDs used in the multispectral imaging system

No	LED Part Number	Color	Dominant Wavelength (nm)	Bandwidth (nm)
1	VLMU3100	Ultra-Violet	405	15
2	SM0603BWC	Blue	428	20
3	SM1204PGC	Green	505	10
- 4	597-3209-202F	Orange	590	20
5	5975112402F	Diffused Red	660	15
6	QBHP684-IR4BU	Near Infra-Red	740	15
7	VSMG3700-GS08CT-ND	Near Infra-Red	850	15
8	751-1253-1-ND	Near Infra-Red	890	20
9	751-1258-1-ND	Near Infra-Red	950	20
10	VSMY3940X01-GS08CT-ND	Near Infra-Red	960	10

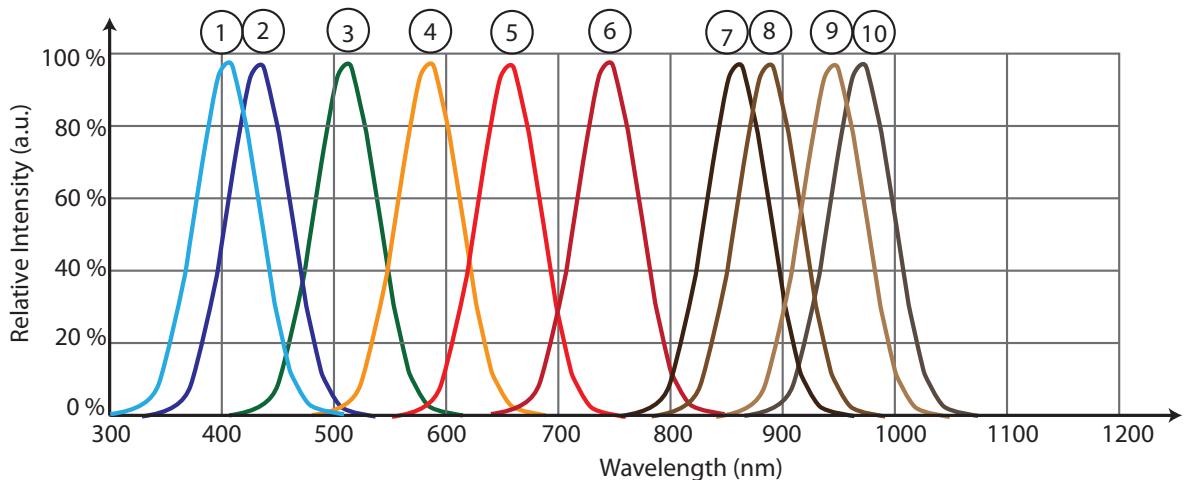


Fig. 4.1 Power spectral distribution for the LEDs on multispectral imaging device. The color of the curve corresponds to the color of the dominant wavelength. Black and brown colors were used for wavelengths at near infra-red region(NIR)

a circumference of a circle (see Fig. 4.2a) and an integrating hemisphere[19] (see Fig. 4.2b) was used to project a uniform light beam.

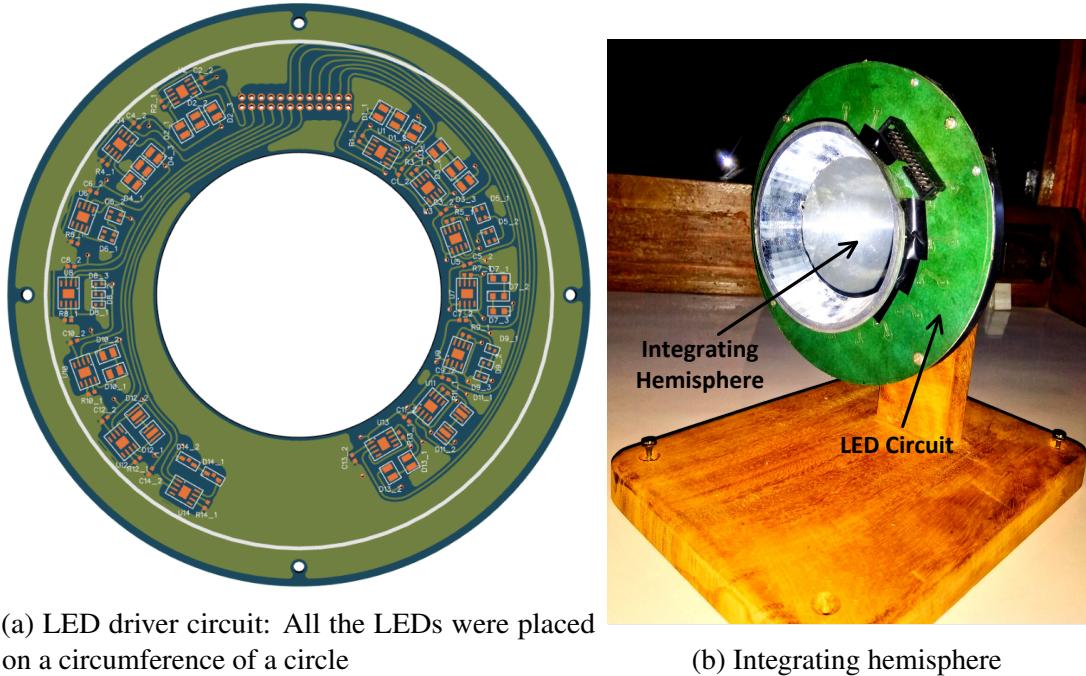


Fig. 4.2 Integrating hemisphere and LED driver circuit

4.2 Capturing RGB Images of Each Wavelength Band

In order to capture a multispectral image, it is required to have a better synchronization between the LED driver circuit and the smartphone camera. For that, a discovery board (STM32F0) and a windows batch script were used. When the batch script was started initially, it automatically turns the smartphone camera on by giving a command via a USB interface between the smartphone and the computer (see Fig. 4.3). When the smartphone camera is turned on, the batch script sends a command to the STM32F0 discovery board to turn on the first LED through a UART interface between the discovery board and the computer (see Fig. 4.3). Once that control signal was received by the discovery board, it turned the first LED in the LED driver circuit on by using a control signal going to the corresponding LED driver IC. After that, the batch script sends a command to the smartphone camera to capture the image which corresponds to the specific wavelength. When all the above were completed, the batch script sends a command to the discovery board to switch to the next LED and this was repeated to all the LEDs. Finally, six RGB images corresponding to six different wavelengths were copied from the smartphone storage to a specified directory in the computer.

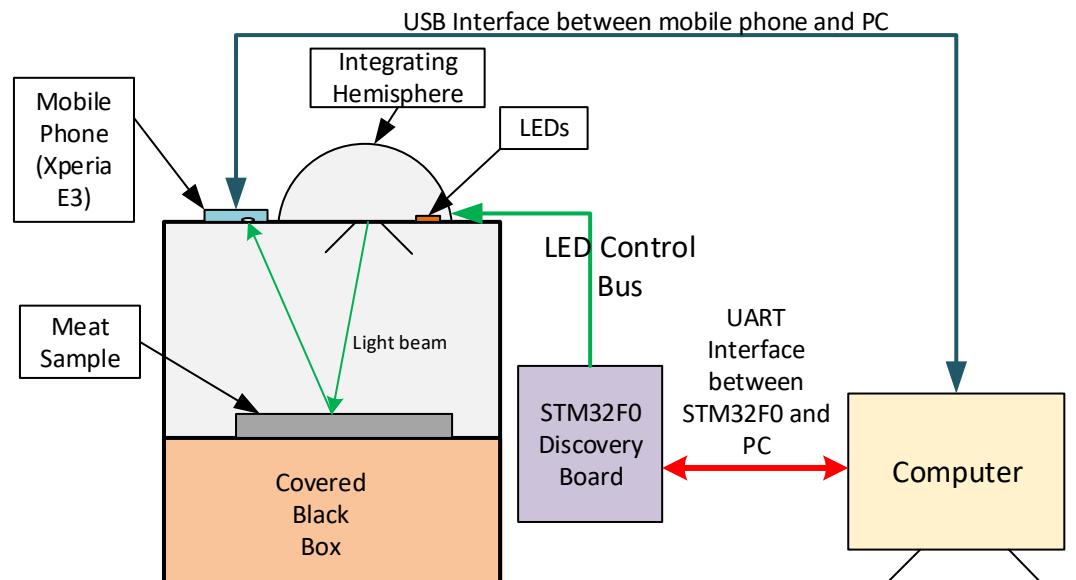


Fig. 4.3 Implemented Multispectral Imaging System as a Block Diagram

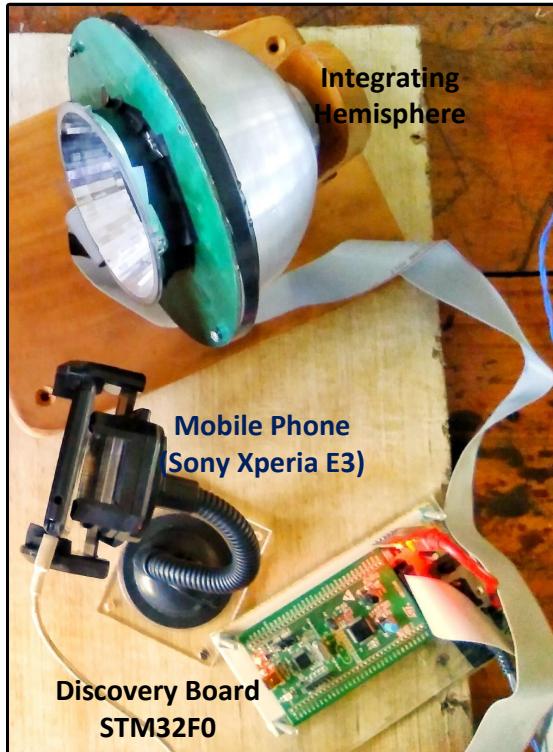


Fig. 4.4 All the parts of the multispectral imaging system - the integrating hemisphere, smartphone camera and discovery board (STM32F0)

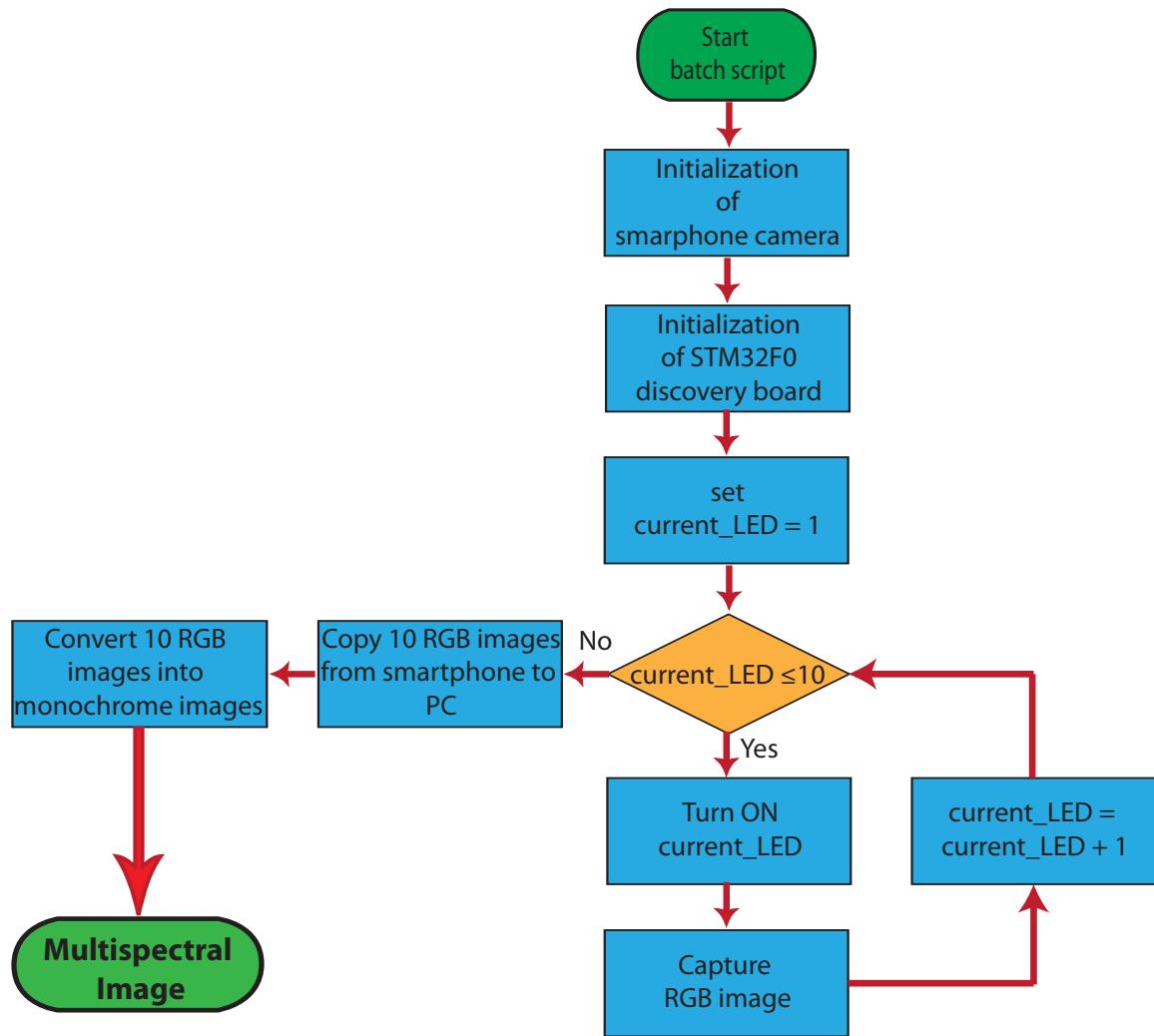


Fig. 4.5 The complete procedure of acquiring a multispectral image by Windows batch script as a flow chart

4.3 Generation of Multispectral Image

After the RGB images of six wavelength bands (actually ten) were captured by a the smart-phone camera, the next step was to generate the multi-spectral image from it as shown in Fig. 4.6. For that, each RGB image (corresponding to specific wavelength band) is required to be converted to a monochrome image. There are various RGB to gray-scale color conversion methods available in the literature. One is by performing a weighted sum of Red (R), Green (G) and Blue (B) components. For an example, MATLAB® uses the following formula to convert an RGB image into a gray-scale image. These particular coefficients were selected by a careful analysis of the human vision system sensitive to red, blue and green colors. Since the human vision system is more sensitive to green, it has the highest coefficient value as given by (4.1).

$$\text{Gray value} = (0.2989R) + (0.5870G) + (0.1140B) \quad (4.1)$$

However, in our Imaging system, we were required to illuminate the object with a specific wavelength band one at a time. When the spectral bands are moving from Ultra-Violet to Near Infra-Red region, the spectral sensitivities of Red, Green and Blue sensors are going to vary accordingly. As an example when the object is illuminated with a wavelength of 430 nm (blue) the blue values of each pixel is more dominant than the Red and Green values. If we are to use above formula to convert that RGB image into the gray-scale image then we would be attenuating the blue values compared to the red and green values which is quite purposeless. Because of this reason, the conversion given by,

$$\text{Gray value} = \sqrt{R^2 + G^2 + B^2} \quad (4.2)$$

was used. With this equation, we will be getting the gray-scale value which is closer to the highest value of R, G and B values. Instead of taking the second norm of the Red, Blue and Green values of each pixel it is also suitable to go for an infinite norm (i.e. maximum value of R,G and B values of each pixel were taken), which will effectively find the maximum sensor response out from R, G and B values.

4.4 Conversion to Reflectance Image

Since the multi-spectral imaging uses spectral and spatial information of an object it is required to obtain spectral information from a set of gray-scale images which typically reflects the light intensity in different spectral bands. The spectral information can be

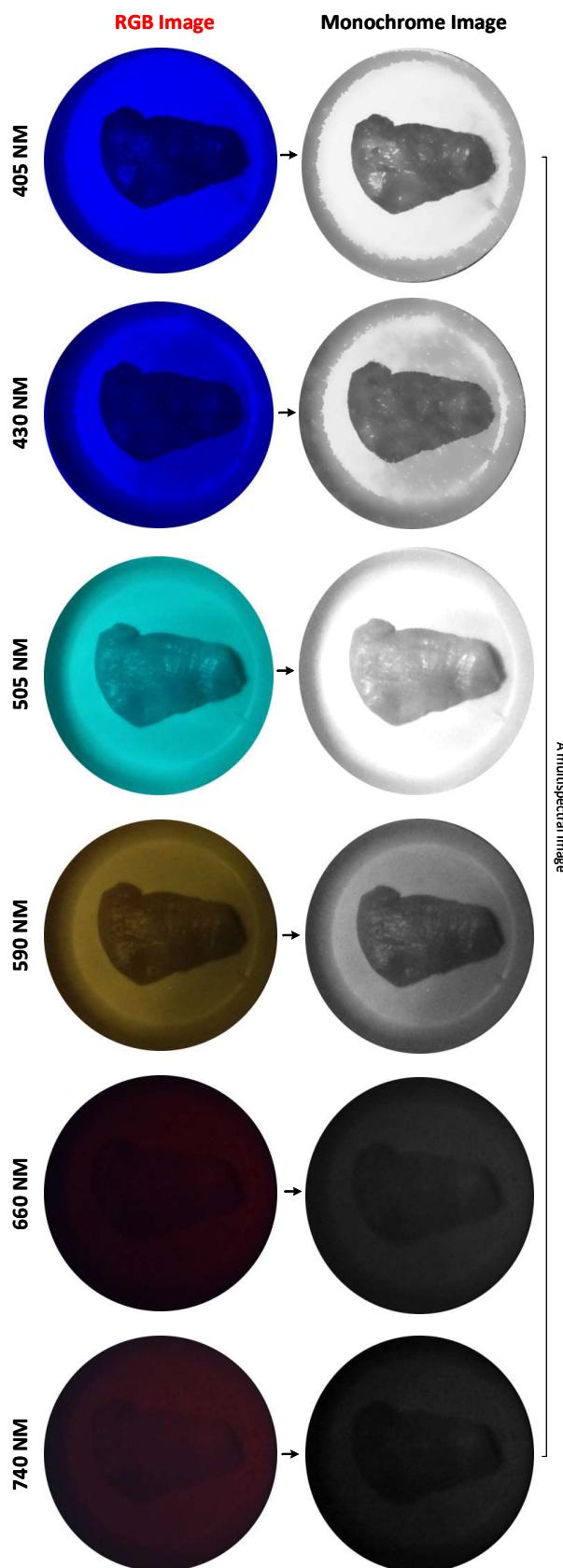


Fig. 4.6 Sample multi-spectral image of a chicken meat sample captured using implemented multispectral imaging system

obtained from reflectance mode, transmission mode or fluorescence mode. Since it is convenient and many research work is done on multi-spectral imaging involving reflectance mode, we have investigated a way to estimate the reflectance information from the gray-scale image. In order to do that multi spectral images of a high reflective ceramic plate were obtained while placing it at the same place where the original sample was placed. In order to avoid the effects coming from the surface non-uniformities of the ceramic plate, several (10) number of multi spectral images were captured while rotating it around its center axis. Then an averaged gray-scale image was taken as the WHITE reference. Similarly, the dark response was acquired by using a plate of lamp soot having a higher absorbent rate and an average gray-scale image was taken as the BLACK reference (see Fig 4.7). Then the transformation[20, 21],

$$\text{Reflectance} = \frac{\text{Pixel Value} - \text{BLACK}}{\text{WHITE} - \text{BLACK}} \quad (4.3)$$

was utilized to obtain the reflectance information by applying it to all the pixels in the gray-scale images corresponding to the six different spectral bands.

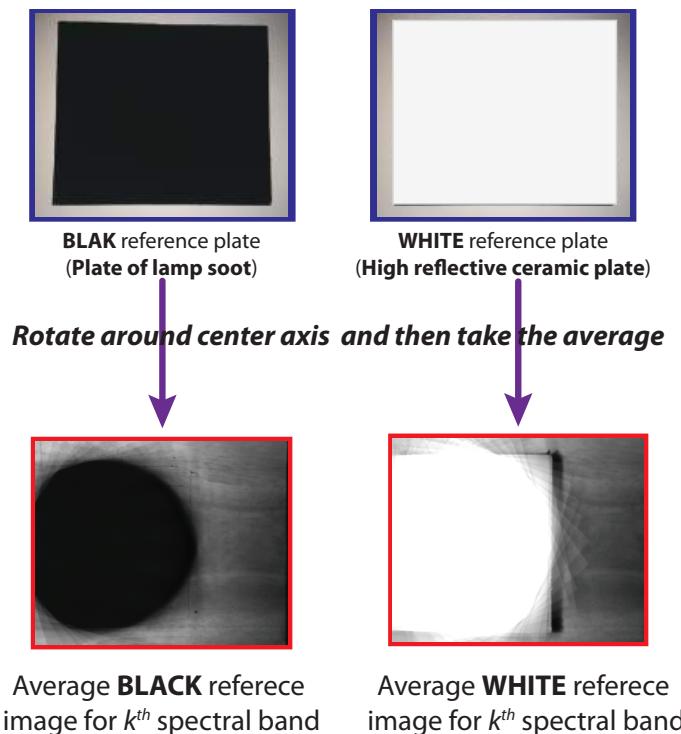


Fig. 4.7 BLACK and WHITE reference plates used to convert intensity image into reflectance image

Chapter 5

Meat Quality Assessment Using The Multispectral Imaging System

Multispectral imaging uses reflectance information of a number of discrete spectral bands to classify samples according to their quality defined using standard parameters. A multispectral image is rich in information compared to a normal RGB image. Therefore, a multispectral image can be used to classify samples more accurately than an RGB image. In this project we have designed a multispectral imaging system that can be used to assess the quality of meat. The system is comprised of six LEDs with nominal wavelengths between 405 nm and 740 nm. The light emitted from LEDs reach the meat sample placed inside a dark chamber through an integrating hemisphere. LEDs are lighted one at a time and images of the meat sample are captured for each flash separately using a smartphone camera. Eventually, all the images of the meat sample, taken at a specific time instance were integrated to form the multispectral image. The meat samples stored at 4 °C were imaged up to four days at predetermined time intervals using the designed system. Once the data acquisition was completed, all the pixels of the multispectral image were represented as points in high dimensional space, which was then reduced to a lower dimensional space using Principal Component Analysis (PCA). It was observed that images of meat sample obtained at different time instances clustered into different regions in the lower dimensional space. The experiment was performed with chicken meat samples. This proves the viability of using multispectral imaging as a non-invasive and non-destructive method of assessing meat quality according to certain quality parameters. Off-the-shelf electronic components and a regular smartphone were used to build the system, thus making the system cost-effective.

5.1 Introduction

Assessment of food quality is a major concern in today's food industry because of public health concerns and customer satisfaction. Consumers are more and more concerned about the safety of the food they purchase in the modern society. Many of the food safety breaches are reported on a daily basis, some of which lead to food poisoning. In this context, determination of meat, one of the most vulnerable food for microbial contamination is of paramount importance. However, most of the methods available for determination of the quality of meat is highly laborious and time-consuming[22].

Generally, meat quality can be defined in terms of texture, flavor, and food safety, which includes implications of both compositional and microbiological properties[23–25]. The visual appearance, textural patterns, and color of fresh meat are the main criteria used by the customers when choosing high-quality meat products and these parameters have a correlation with the properties such as chemical composition and microbiological quality[25]. The conventional methods of determining meat quality evaluation are time-consuming, invasive, destructive and need human inspection. Therefore evaluation of meat quality in today's meat processing lines as well as in markets require fast, accurate and cost-effective assessment method which overcomes all the disadvantages of traditional methodology. In this chapter, we present a way to assess the quality of a meat sample in real time with high accuracy using a low-cost multispectral imaging system which uses a smartphone RGB camera (SONY Xperia E3-D2203 rear camera with the resolution of 480×854 pixels) as the sensing device.

5.1.1 Advantages of Using Multispectral Imaging Technology in Food Analysis

- No sample preparation is required.
- Chemically free assessment method.
- Once the calibration model is build and validated, it becomes an extremely simple analysis method.
- It is a non-invasive and non-destructive method, so that the same sample could be used for other purposes and analyses.
- More economical compared with traditional methods. Owing to saving in labor and time.

- Flexibility in choosing any Region of Interest (ROI) in the image even after the image acquisition.
- Incomparable for process monitoring and real-time inspection.
- It is able to determine several constituents simultaneously in the same sample.

5.2 Methodology

The meat samples stored at 4 °C were imaged up to four days at predetermined time intervals using the designed system. At the same time the microbial count at the meat samples, pH value and other organoleptic properties such as color, odor, overall appearance and the surface texture (sliminess) were evaluated and recorded. Once the data acquisition was completed, all the pixels of the multispectral image were represented as points in high dimensional space, which was then reduced to a lower dimensional space using Principal Component Analysis (PCA). It was observed that images of meat sample obtained at different time instances clustered into different regions in the lower dimensional space. The experiment was performed with chicken meat samples. This proves the viability of using multispectral imaging as a non-invasive and non-destructive method of assessing meat quality according to certain quality parameters. The key steps involved with meat/food quality analysis using the developed multispectral imaging system is shown in Fig. 5.1.

5.2.1 Measurement of Meat Quality Parameters

Fresh skinless and boneless chicken breast fillets ($3\text{ cm} \times 2\text{ cm}$) were obtained from the Livestock field station at Mawela, Uda Peradeniya. They were packed in polythene packages (see Fig. 5.2 (a)) and were transported to the laboratory immediately in a cool box and were stored at -18 °C until they were used for the microbial analysis. Meat samples were brought to 4 °C and were stored in a refrigerator (see Fig. 5.2 (b)). Samples were drawn after 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs, and were macerated immediately in a bag mixer (CC, BagMixer®400, France) under near sterile conditions (see Fig. 5.2 (c)(d)) and the resulting slurry was used for analysis. The total plate count (see Fig. 5.2 (e)), pH (see Fig. 5.2 (f)) and organoleptic properties in meat samples were determined in triplicate. For the determination of the Total Plate Count (TPC), each meat sample was homogenized separately in 9 ml buffered peptone water (1.0 g peptone, 8.5 g NaCl, 1000 ml distilled water) in a stomacher for 1 minute and a dilution series were prepared by adding 1 ml of suspension in 9 ml buffered peptone (see 5.3). Subsequently, 1 ml aliquots

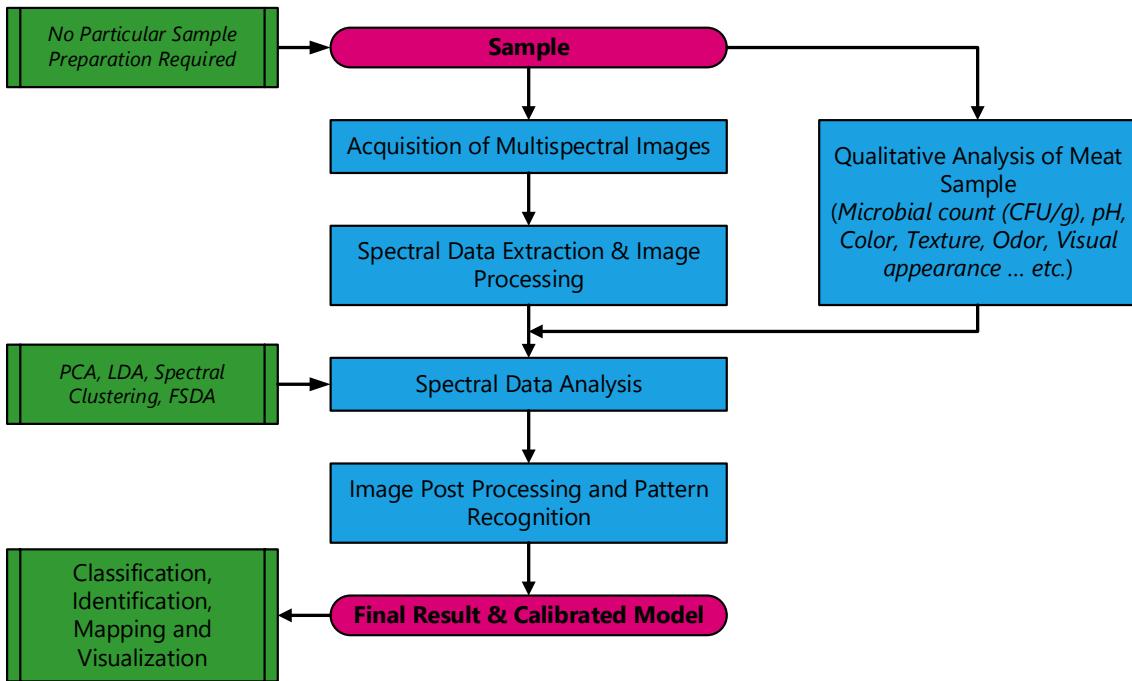


Fig. 5.1 Flowchart of the key steps involved in meat quality analysis using the multispectral imaging system

of each dilution was placed onto a 3M™ Petrifilm™ Aerobic Count Plate and incubated at 30°C for 48 hrs in an incubator (IO 600, Yamato, Japan)[26]. At the end of the incubation period, all the red color colonies were counted using a colony counter(DC-3, Kagayaki Irika Kogyo, Japan) and the results were expressed as CFU/g as in using the equation[27],

$$\text{Colony Forming Units (CFU)}/g = \frac{\sum C}{(n_1 + 0.1n_2) \times d} \quad (5.1)$$

where,

$\sum C$ = Sum of colony count in all Petri films used for calculation,

n_1 = No. of Petri films used in first dilution,

n_2 = No. of Petri films used in second dilution,

d = Dilution factor of first dilution used in calculation.

Simultaneously, pH value of the sample was determined using a laboratory pH meter (ECPH27005E, Eutech, Singapore) and organoleptic properties such as color, odor, overall

appearance and the surface texture (sliminess) were evaluated and recorded as shown in Tab. 5.1. The microbial growth curve plotted on a semi-log, as shown in Fig. 5.4, gives approximately a linear relationship with storage time.

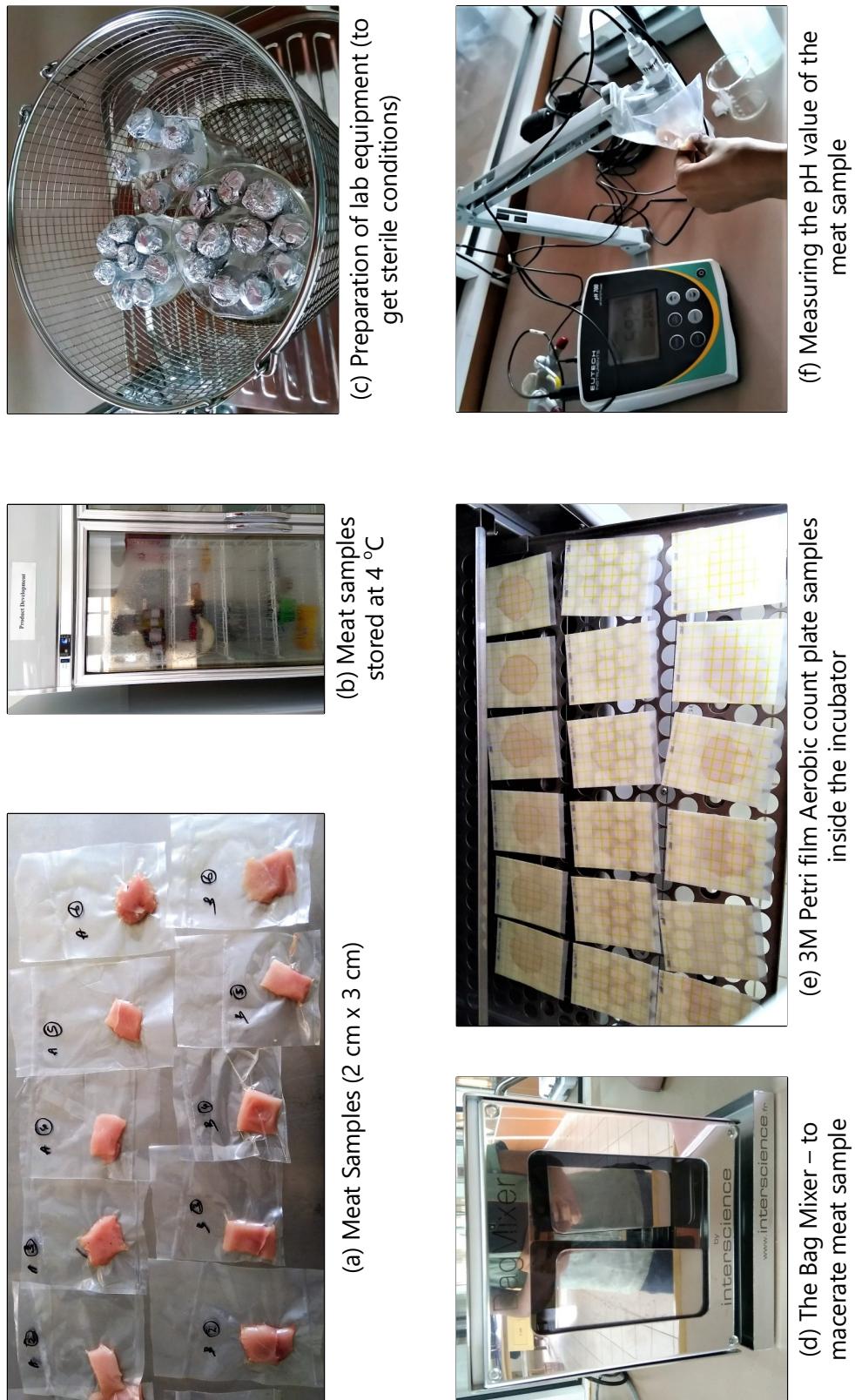


Fig. 5.2 The key steps involved with measuring microbial count (CFU/g) and pH value in a lab environment

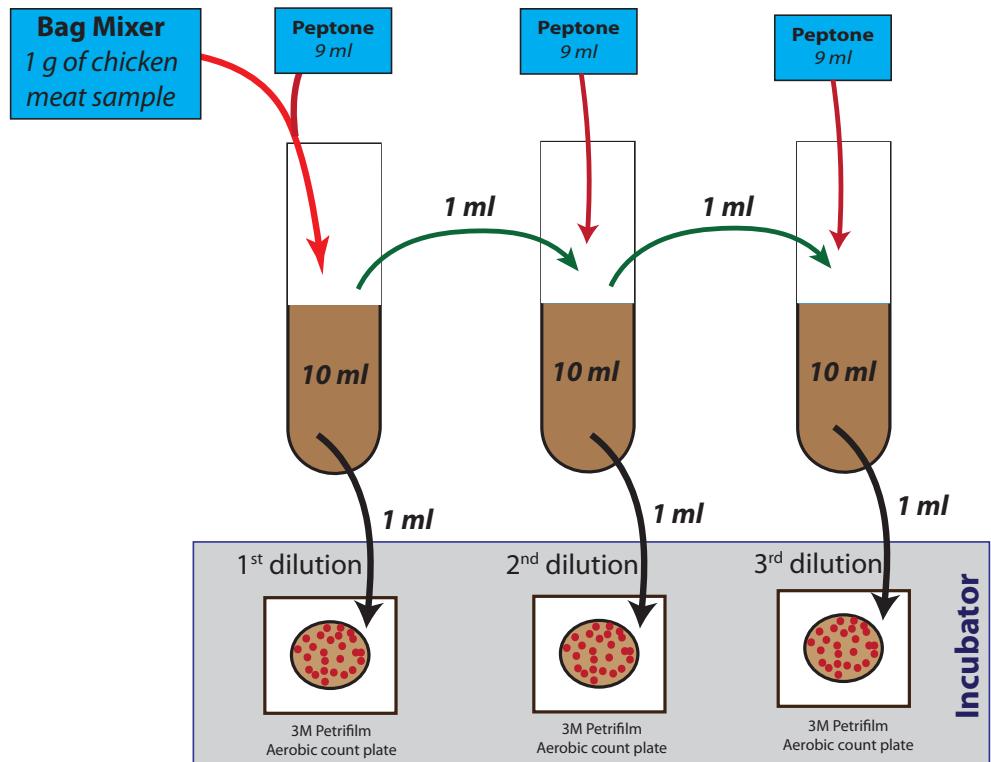


Fig. 5.3 The method of determining microbial count (Colony Forming Units (CFU) /g with the aid of dilution series.

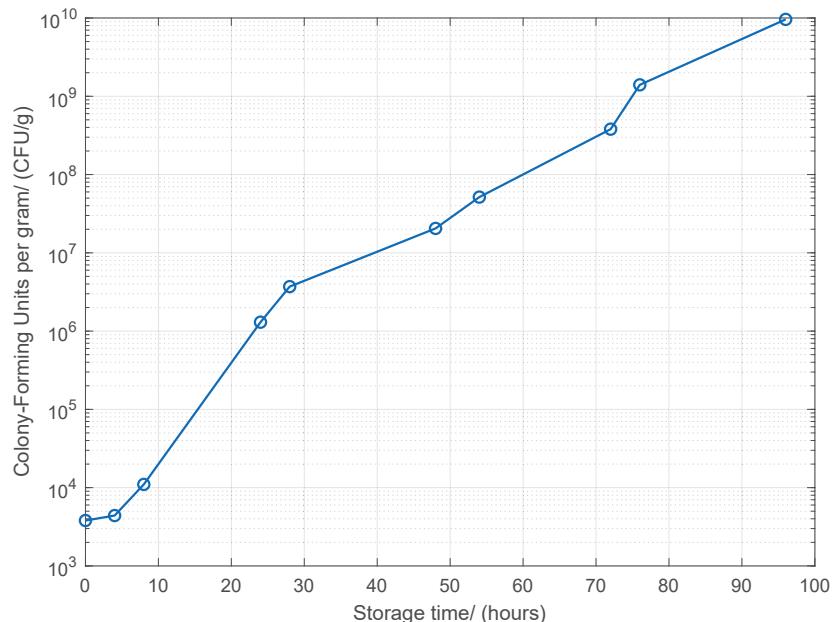


Fig. 5.4 Variation of microbial count (Colony Forming Units per gram - CFU/g) for storage time of 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs.

Table 5.1 Meat Quality Parameters for Different Storage Times at 4 °C

Storage Time at 4.00 °C (hrs)	Average pH Value	Microbial Count (CFU/g)	Color	Odor	Sliminess
0	6.39 (at 27.7 °C)	3.8×10^3	Reddish pink	Normal	absent
4	5.90 (at 28.0 °C)	4.4×10^3	Pink	Normal	absent
8	5.83 (at 26.1 °C)	1.1×10^4	Pink	Normal	absent
24	5.92 (at 26.2 °C)	1.3×10^6	Pink	Rotten	absent
28	5.80 (at 28.3 °C)	3.7×10^6	Pink	Rotten	absent
48	5.89 (at 32.1 °C)	2.1×10^7	Pink	Rotten	present
54	6.11 (at 29.1 °C)	5.2×10^7	Pink	Rotten	present
72	5.80 (at 29.0 °C)	3.8×10^8	Pink	Rotten	present
76	6.05 (at 29.7 °C)	1.4×10^9	Whitish pink	Rotten	present
96	6.02 (at 29.2 °C)	9.6×10^9	Whitish pink	Rotten	present

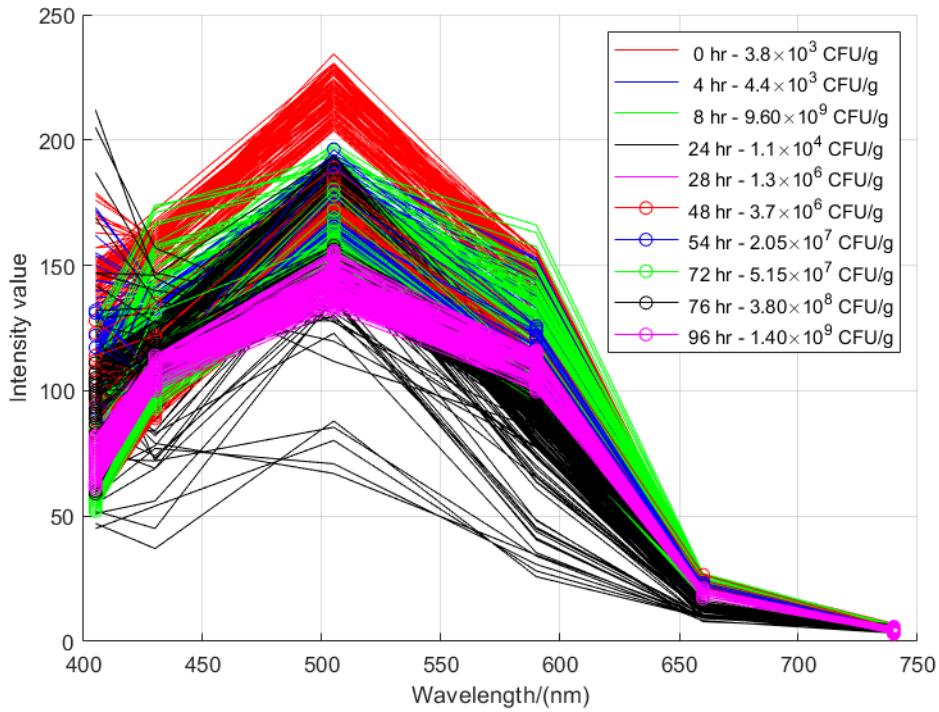


Fig. 5.5 Spectral signatures of all the pixels for storage time of 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs. Microbial count for each storage time also indicated in the legend.

5.3 Algorithm and Results

The multispectral images of chicken samples corresponding to the storage times, 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs were loaded into MATLAB®. Then for the analysis, 121 pixels from each of the multispectral image were selected randomly on the area of the meat piece. Next, the spectral intensities of those pixels were arranged in a matrix of 121×6 for each of the storage time. Here, the 121 rows in the matrix correspond to the 121 pixels selected from the meat sample and the 6 columns correspond to the number of wavelengths of the imaging system (405 nm, 430 nm, 505 nm, 590 nm, 660 nm and 740 nm). The same procedure was repeated for rest of the multispectral images with different storage times. Once all the matrices corresponding to different storage times were made, the analysis part was carried out.

First, to understand the nature of the multispectral image dataset, intensity values of all the pixels were plotted against the spectral bands (called the spectral signature) for different storage times as shown in Fig. 5.5. Then, the average spectral signature curves for different storage times were obtained as shown in Fig. 5.6. Since it was difficult to identify the

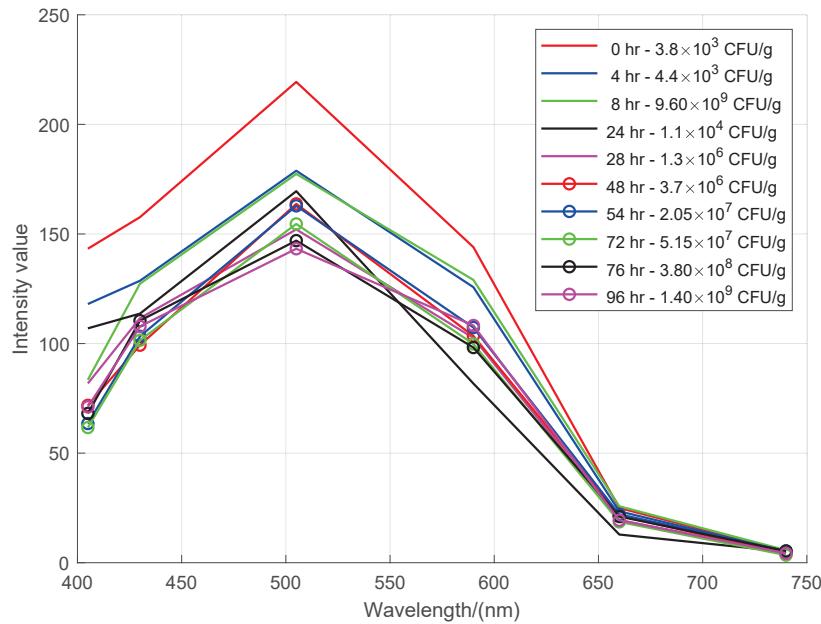


Fig. 5.6 Average spectral signatures of chicken meat samples for storage time of 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs. Microbial count for each storage time also indicated in the legend.

relationship between the spectral signatures and the quality (i.e. Microbial count) of the meat samples, an algorithm known as the Principal Component Analysis (PCA) was performed on the dataset. The PCA is a dimensionality reduction technique on which original data set is projected into a new set of axes known as principal components. The projected dataset along the first principal component axis (PC1) has the highest variance, while the projected dataset along the second principal component axis (PC2) has the second highest variance and so on. The PCA equations are given by,

$$Cov = E\{(X_i - \mu)(X_i - \mu)^T\} \quad (5.2)$$

$$Cov = \frac{1}{N} \sum_{i=1}^N (X_i - \mu)(X_i - \mu)^T \quad (5.3)$$

$$Cov V = \lambda V \quad (5.4)$$

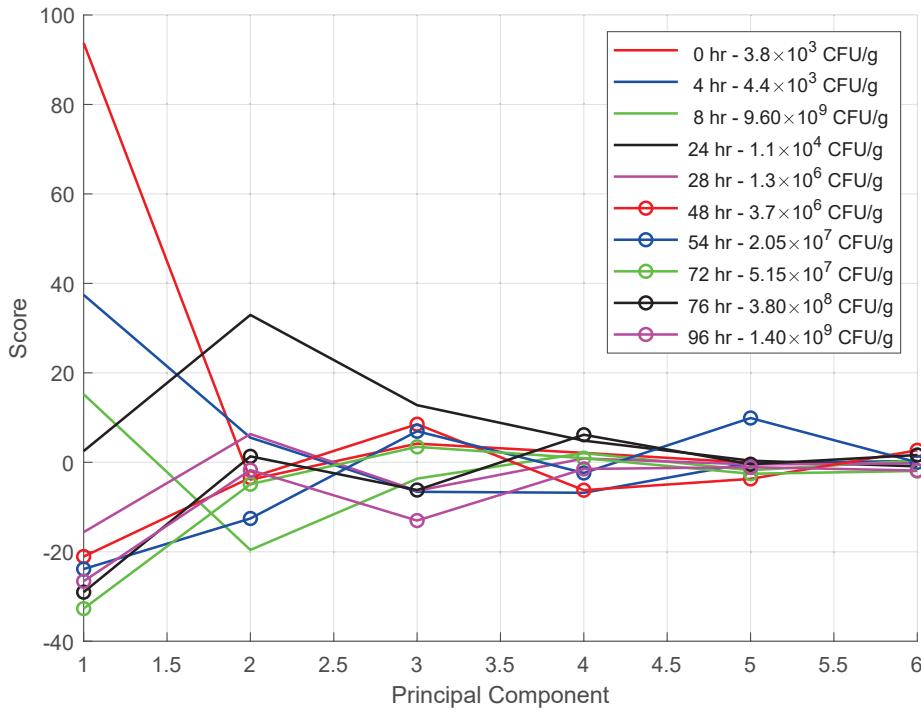


Fig. 5.7 Average signature curves for storage time of 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs after performing the Principal Component Analysis (PCA) on the original data set. Now, x-axis represents the six principal components and y axis represents the average score for each storage time.

where,

Cov = Covariance matrix (6×6),

E = Expectation operator,

X_i = Vector representation of i^{th} pixel (6×1),

μ = Vector representation of the mean spectral intensity
of all pixels (6×1),

N = Total number of pixels ($121 \times 10 = 1210$),

V = Eigenvector of the covariance matrix (6×1),

λ = Eigenvalue of the covariance matrix.

Then, average value of all the pixels along each principal component axis was plotted for different storage times as shown in Fig.5.7.

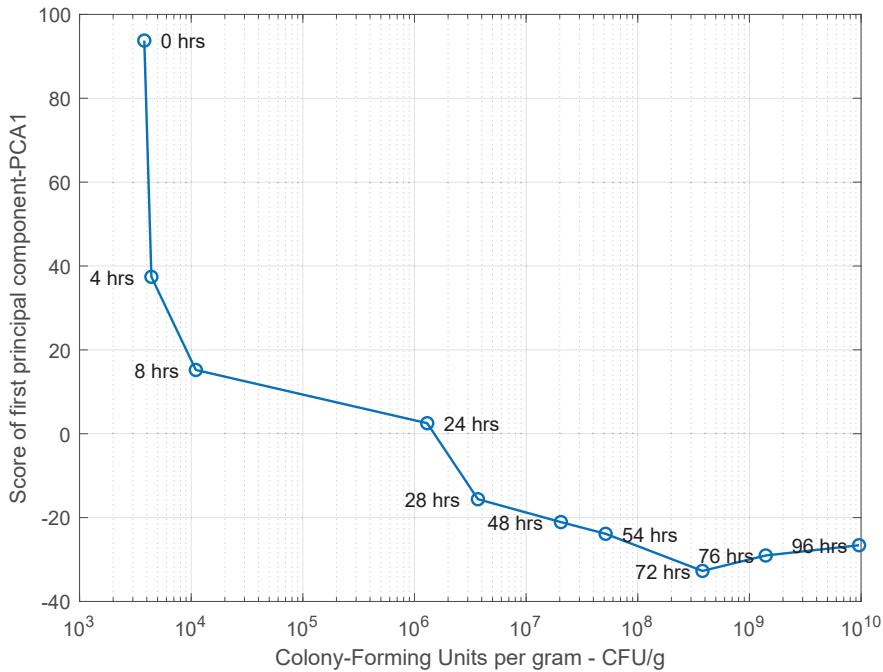


Fig. 5.8 Average value of first principal component (PC1) for storage time of 0 hrs, 4 hrs, 8 hrs, 24 hrs, 28 hrs, 48 hrs, 54 hrs, 72 hrs, 76 hrs and 96 hrs.

The average values of first principal component (PC1 value) were plotted with colony-forming units per gram (CFU/g) as shown in Fig. 5.8. Therefore, this relationship between average score of first principal component (PC1) and CFU/g value can be used to estimate the quality (CFU/g value) of an unknown chicken meat sample. Finally, all the pixels were plotted in the reduced dimensional space as shown in Fig. 5.9 and it was observed that pixels are crusted together according to the storage time.

5.4 Conclusion

In this paper, we presented a smartphone RGB camera based multispectral imaging system that can be used to assess the microbial quality of chicken and thereby to determine the safe storage time period. The number of Colony Forming Units in one gram of meat sample (CFU/g) was taken as the quality of a meat sample while measuring other quality parameters such as pH value, color, smell, odor, and presence of slime simultaneously. Then it was observed that CFU per gram (CFU/g) has a strong correlation with spectral parameters (average score of PC1) as shown in Fig. 5.8. In order to clearly visualize the relationship between CFU/g with spectral parameters, the Principal Component Analysis was used. Hence,

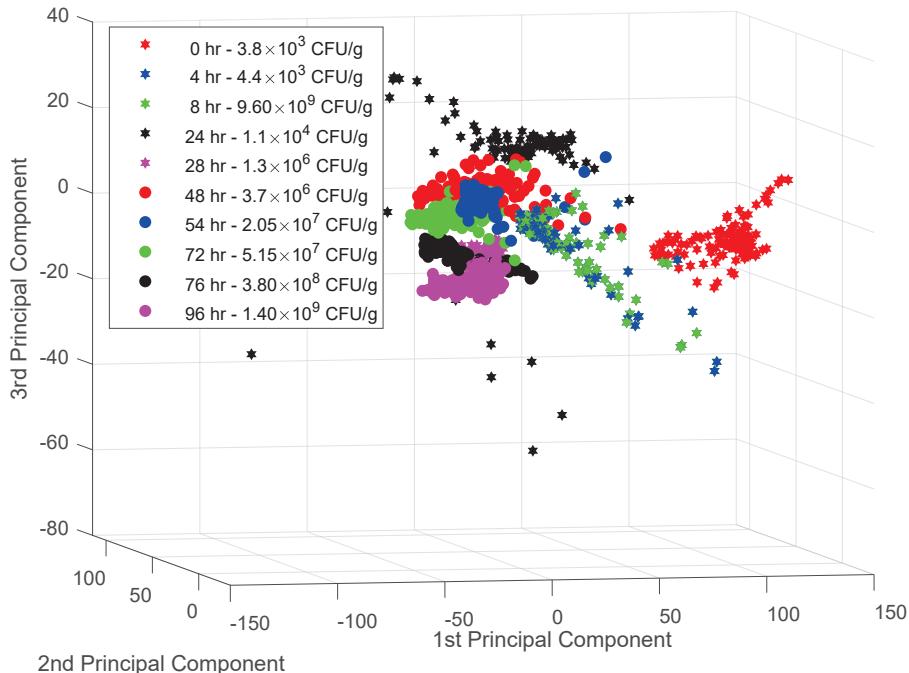


Fig. 5.9 Scatter plot of all the pixels in reduced dimensional space - PC1, PC2 and PC3. In this case pixels are clustered together according to the storage time.

by using smartphone RGB camera as the sensing device of multispectral imaging system we were able to estimate the quality of a meat sample in a non-invasive and non-destructive manner while minimizing chemical cost, labor cost and measuring time compared with traditional meat quality assessment methods.

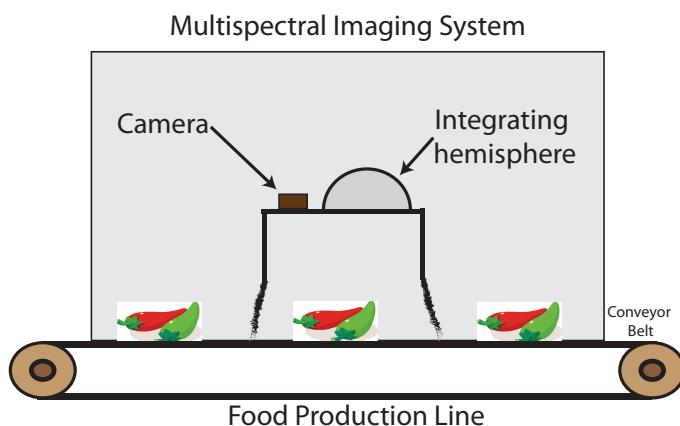


Fig. 5.10 The multispectral imaging system installed in a food processing line to automate quality assessment

Chapter 6

Validation of multispectral imaging for the detection of adulterants in turmeric samples

Turmeric (*Curcuma longa*) has been used over a hundred of years as a spice and as an antiseptic in traditional medicine for healing a variety of ailments. Due to its high popularity and demand in the international market, turmeric powder is often adulterated. Different types of flour such as wheat flour and rice flour mixed with colorants are the main adulterants regularly identified in turmeric powder. The use of unpermitted dyes such as fabric dyes as colorants potentially poses serious health risks as some of them are carcinogenic. In this backdrop, the identification of adulterated turmeric is of paramount importance. However, the existing methods are laborious, expensive and time-consuming. In this study, we explore the possibility of adopting multispectral imaging for detecting the presence of wheat flour and tartrazine in turmeric powder. A multispectral imaging system was developed with the dominant wavelengths ranging from 405 nm to 950 nm with a resolution of nine spectral bands. A mathematical algorithm was developed based mainly on Principal Component Analysis and Bhattacharyya Distance. The relationship between the Bhattacharyya Distance and the adulteration level was modeled as a second-order polynomial with an $R^2 = 0.9911$ and the results were validated using an independent set of samples with an $R^2 = 0.9816$.

6.1 Introduction

Turmeric (*Curcuma longa*) which is used all over the world as a spice and an antimicrobial, has been used in traditional medicine for thousands of years. Curcumin (diferuloylmethane)[28]

present in turmeric has been used extensively in Ayurvedic medicine for centuries because of its proven bioactivities such as its antioxidant[29], analgesic, anti-inflammatory[30], antiseptic [31], and anticarcinogenic qualities [32, 33]. Hence, there is a growing demand for turmeric in international trade because of its high popularity. Due to these reasons, turmeric powder has been often adulterated in order to boost profits [34, 35].

Materials such as rice flour, wheat flour, starch and chalk dust stained with different types of dyes and colorants are commonly added to turmeric powder. Among these colorants, tartrazine (synthetic lemon yellow azo dye, E number: 102, C16H9N4Na3O9S2), metanil yellow (yellow azo dye: C18H14N3NaO3S) and lead chromate ($PbCrO_4$) are the most commonly used colors as reported by researchers [36–38]. There have been instances where fabric and other industrial dyes and stains have also been used to adulterate turmeric powder.

A variety of conventional methods have been effectively used to detect adulterants present in turmeric powder such as rapid color test[38], microscopic analysis[39] and thin layer chromatography[40]. Similarly, other methods such as high-performance liquid chromatography[41, 42], micellar chromatographic method, and high-performance capillary electrolysis have been used to determine the presence of metanil yellow and aniline dyes. Despite their high accuracies and satisfactory detection limits, these conventional methods are limited as a field deployable technology due to their operational complexity, destructive nature, extensive sample preparation, requirement of chemicals, and difficulty in automating the detection process. Moreover, most of these chemical analyses focus on the identification of the presence of a limited number of stains or colorants.

In this light, the adoption of optical detection methods are gaining growing popularity for the quality evaluation of foods as they can overcome some of the limitations of conventional methods. A variety of imaging systems and screening techniques have been proposed in the literature for detecting the presence of adulterants of various foods such as meat[43], honey[44], fruits and vegetables[3, 45], powders[38, 46], etc.

This study involves the use of ultra-violet (UV), visible and near infra-red (NIR) multispectral imaging for determining the presence of tartrazine, wheat flour, and rice flour in turmeric powder. A multispectral imaging system was developed [6, 47] based on the selected wavelengths from UV, visible and NIR regions of the electromagnetic spectrum with the resolution of nine spectral bands. These type of multispectral imaging systems have been used for the assessment of the quality of foods such as meat, fruits, and vegetables [48, 49, 3, 18, 45]. However, they are not widely used for the detection of food adulterants. More complex spectroscopic methods are found in the literature; however, with the application of appropriate signal processing knowledge together with a few spectral bands, a good curve fitting ($R^2 = 0.9912$) and validation accuracy ($R^2 = 0.9816$) was achieved. A

set of narrow-band LEDs were used in the multispectral imaging system, thereby reducing the hardware cost and complexity of the system. Then, the image processing knowledge was applied appropriately to determine the adulteration level. In most of the literature, PCA algorithm together with a regression model has been generally used to develop the system model [50–53]. In this study, it has been proposed how the use of the Bhattacharya distance, the Jeffries–Matusita (JM) distance, and the Maximum likelihood estimation can be utilized effectively to accurately predict the adulteration level followed by an image pre-processing step and the Principal Component Analysis (PCA).

Thus, the objective of this study was to evaluate the potential of multispectral imaging in tandem with data analytic techniques such as the PCA, the Bhattacharyya distance and the JM distance to quantify the presence of tartrazine, wheat flour, and rice flour in the turmeric powder.

6.2 Materials and Methods

6.2.1 Preparation of samples

Authentic turmeric (*Curcuma longa*) powder samples were prepared using fresh turmeric rhizomes collected from New Ranweli Spice Garden, Kandy, Sri Lanka. The cleaned rhizomes were sun-dried, finely ground and passed through a sieve of 300 µm mesh to obtain the powder with uniform particle size. Adulteration mixture was prepared by mixing wheat flour with tartrazine (synthetic yellow azo dye, E number - E102) at 9:1 (w/w) ratio. The authentic turmeric sample was adulterated with the prepared mixture by homogeneously mixing the adulterant at different ratios ranging from 0% to 40% (w/w). Thirty replicates were prepared for each adulteration level of 0%, 5%(w/w), 10%(w/w), 15%(w/w), 20%(w/w), 25%(w/w), 30%(w/w), 35%(w/w) and 40%(w/w). These samples (totally $9 \times 30 = 270$ samples) were used for the purpose of calibration. At the same time, sixty more independent samples were prepared with known adulteration levels from 0% to 40% (w/w) and those samples were used for validation. Moreover, fifteen different commercially available turmeric powder samples were collected from three different geographical regions and five replicates were taken from each sample for the multispectral image analysis. Finally, the multispectral images of the calibration samples, validation samples and commercially available samples were captured using the developed multispectral imaging system.

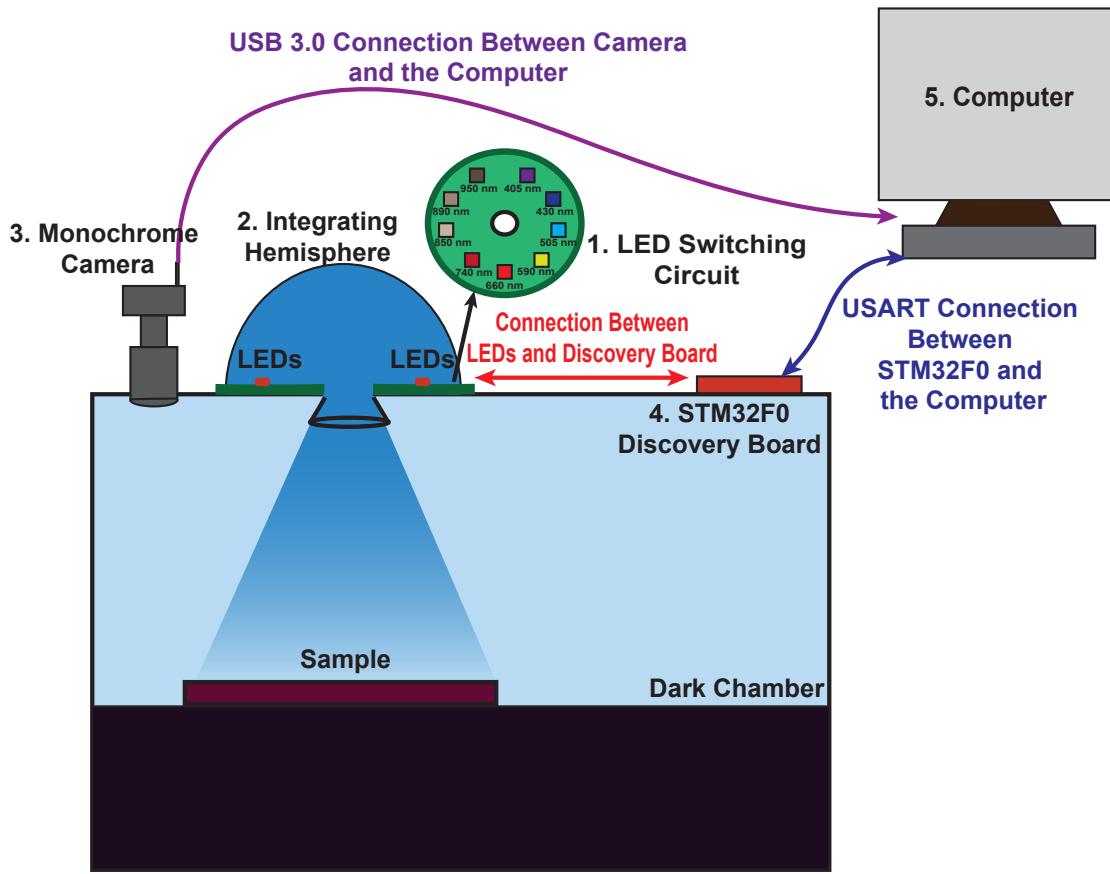


Fig. 6.1 Schematic of the developed multispectral imaging system

6.2.2 Multispectral Image acquisition

A multispectral imaging system [49, 54] was developed based on the wavelengths selected from Ultra-Violet (UV) to Near Infra-Red (NIR) region of the electromagnetic spectrum. As shown in Figure 6.1, the system consists of five components: a LED switching circuit (LEDs with dominant wavelengths 405 nm, 430 nm, 505 nm, 590 nm, 660 nm, 740 nm, 850 nm, 890 nm and 950 nm), an integrating hemisphere (Radiometry, 2009)(inner diameter – 130mm and made up of Al), a monochrome camera (FLIR Blackfly S Mono, 1.3 MP, USB3 Vision camera, Resolution – 1280 x 1024, ADC – 10 bit), a discovery board (STM32F0DISCOVERY) and a computer (Acer Aspire E 15, Intel core i5 @2.8GHz, 8GB DDR4 Memory). The details of the LEDs used in the LED switching circuit are given in Table 6.1. The intensities of all the LEDs were adjusted to a constant approximated value using the LED driver ICs (MAX16839ASA+). The resulting power spectral density distribution curves of the LEDs are shown in Figure 6.2.

6.3 Multispectral Data Analysis

In most of the multispectral imaging applications, the final goal of the research is targeted towards a specific application, in which researchers first identify the most suitable wavelengths for that application as a priori. Then a multispectral imaging system can be implemented based on the selected wavelengths as described in the previous section. After that, a common way of reducing (if required) the number of images should be performed using a dimensionality reduction technique i.e. Principal Component Analysis (PCA), Fishers/ Linear Discriminant analysis (FDA/LDA) or spectral clustering. Most of the time, results from this process are fed into a machine learning algorithm. In many of the system identification problems, linear regression models have been used to create the prediction model. In this research, we proposed an algorithm based on the Bhattacharyya distance and the JM separability measure was proposed to determine the adulteration level of turmeric powder.

6.3.1 Image Pre-Processing

For a multispectral imaging system, there exist different types of random noise including camera read-out noise, wire connection and data transfer noise between the camera and the computer, electronic noise inherent to the camera such as dark current, and noise from digitizing while doing Analog to Digital (A/D) Conversion [55–57]. These noise values will have a definite impact on the results generated from multispectral image analysis. In the image preprocessing step, dark current subtraction and spectral low pass filtering were performed to remove random noise.

Dark current is mainly due to the currents that are generated as a result of the formation of the depletion region [58] and irregularities of the silicon lattice surface of the photo-diode. This current is added to the current from drift and diffusion in the photo-sensor so that even

Table 6.1 Details of the LEDs used in the LED switching circuit (Purchased from Digi-Key Electronics - <https://www.digikey.com>)

LED No.	Region (UV, Visible and NIR)	Manufacturer Part Number (Manufacture)	Dominant Wavelength (nm)	Bandwidth (nm)
1	UV	VLMU3100 (Vishay)	405	10
2	Visible	SM0603BWC (Bivar)	430	50
3		SM1204PGC (Bivar)	505	20
4		5973209202F-ND (Dialight)	590	10
5		5975112402F (Dialight)	660	20
6	NIR	QBHP684-IR4BU (QT Brightek (QTB))	740	20
7		VSMY2850G (Vishay)	850	10
8		VSMF4710-GS08 (Vishay)	890	10
9		VSMS3700-GS08 (Vishay)	950	20

if there is no current from external light, the pixel will still measure the dark current. So, instead of just using sample data for the multispectral image analysis, dark current can be subtracted from the sample data prior to any further analysis. The dark current images were taken at the beginning of each multispectral image acquisition. Then dark current subtraction was performed on each spectral image using the equation given by [59, 60],

$$P[\lambda] = S[\lambda] - D \quad (6.1)$$

Where $P[\lambda]$ is the dark current removed sample data at wavelength λ , $S[\lambda]$ is the raw sample data at wavelength λ and D is the dark current image captured at the beginning of the multispectral image acquisition. This will remove some inherent noise generated in the image sensor which is better than using the raw sample data as it is for the analysis.

The median filtering and 2D adaptive filtering were performed separately on the dark current subtracted sample data (P) to remove random noise. The median filtering[61] is the simplest way of smoothing the sample data. The median filtering was performed by running a moving average filter on each pixel with a window size of (3×3) using the equation given by,

$$P_{ma}^* = \frac{1}{N} \sum_{k=-m}^{k=+m} \sum_{l=-m}^{l=+m} P[i+k, j+l] \quad (6.2)$$

Where P_{ma}^* is the smoothed sample data after the moving average filtering, P is the dark current subtracted image, $N = (2m + 1) \times (2m + 1)$ is the total number of pixels in the selected window and (i, j) is the location of a pixel in the image.

At the same time, an adaptive Wiener filter was used to remove additive noise in the sample images and compared them with results obtained using the moving average filtering. The main difference in adaptive Wiener filtering compared to the moving average filtering is that filter parameters are space variant. In other words, filter coefficients (pixel mean - $\mu[i, j]$ and variance - $\sigma^2[i, j]$) change as a function of $[i, j]$. Therefore, the pixel mean ($\mu[i, j]$) and the variance ($\sigma^2[i, j]$) are not fixed for the entire image and they are calculated locally at each pixel location. The local mean ($\mu[i, j]$) at the pixel location $[i, j]$ can be calculated using the equation given by,

$$\mu[i, j] = \frac{1}{N} \sum_{k=-m}^{k=+m} \sum_{l=-m}^{l=+m} P[i+k, j+l] \quad (6.3)$$

Where P is the dark current subtracted image and $N = ((2m + 1)^2)$ is the total number of pixels in the window. The local variance $\sigma^2[i, j]$ at the pixel location $[i, j]$ can be evaluated

using the equation given by,

$$\sigma^2[i, j] = \left(\frac{1}{N} \sum_{k=-m}^{k=+m} \sum_{l=-m}^{l=+m} P^2[i+k, j+l] \right) - \mu^2[i, j] \quad (6.4)$$

The white noise variance (v^2) is assumed to be constant over the entire image and can be calculated by averaging all the locally estimated variances $v^2(i, j)$ using the equation given by,

$$v^2 = \frac{1}{L \times W} \sum_{i=0}^{i=L} \sum_{j=0}^{j=W} \sigma^2[i, j] \quad (6.5)$$

Where L is the length of the sample image (No. of pixels), W is the width of the sample image (No. of pixels) and $(L \times W)$ is the total number of pixels in the image. Finally, the processed image (P_w^*) can be obtained using the equation given by [61],

$$P_w^*[i, j] = \mu[i, j] + \frac{\sigma^2[i, j] - v^2}{\sigma^2[i, j]} (P[i, j] - \mu[i, j]) \quad (6.6)$$

6.3.2 Spectral Signatures

An example of a multispectral image of turmeric powder sample which was obtained after the image pre-processing step is shown in Figure 6.3– (a). A random pixel (red color dot) was selected in the region of turmeric powder sample and the grayscale values (between 0 and 255) were plotted against the wavelength to obtain the spectral signature as shown in Figure 6.3 – (b).

For the extraction of mean spectral signature, the multispectral images corresponding to each adulteration level (0%, 5%(w/w), 10%(w/w), 15%(w/w), 20%(w/w), 25%(w/w), 30%(w/w), 35%(w/w) and 40%(w/w)) were represented as two dimensional data matrixes $X_0(3000 \times 9)$, $X_5(3000 \times 9)$, $X_{10}(3000 \times 9)$, $X_{15}(3000 \times 9)$, $X_{20}(3000 \times 9)$, $X_{25}(3000 \times 9)$, $X_{30}(3000 \times 9)$, $X_{35}(3000 \times 9)$ and $X_{40}(3000 \times 9)$, in which columns represent the wavelengths and rows represent the observation points selected from each replicate as shown in Figure 6.4 . Resulting mean spectral signature curves are shown in Figure 5. From the mean spectral signature curves shown in Figure 6.5–(a), it is clear that mean intensity values are increasing with the adulteration level for the wavelengths 405 nm, 430 nm, 505 nm, 590 nm, and 660 nm while no direct relationship was observed for the wavelengths 740 nm, 850 nm, 890 nm and 950 nm.

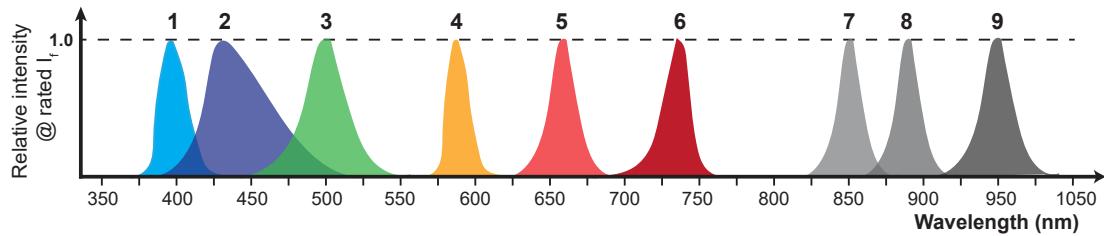


Fig. 6.2 The power spectral density distribution curves of the LEDs used for LED switching circuit

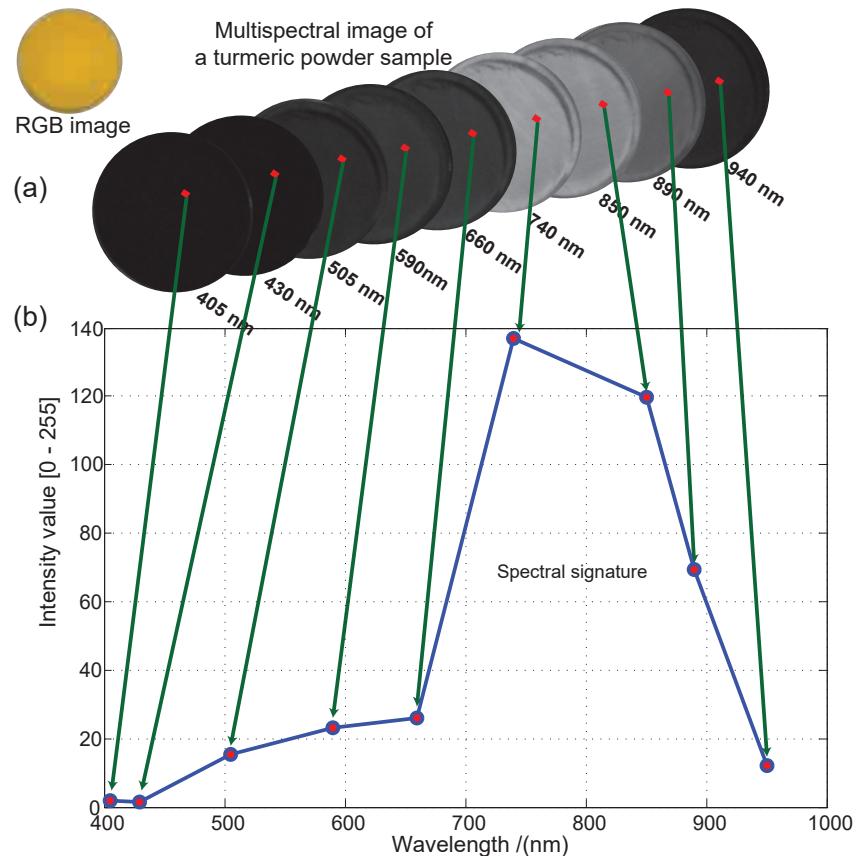


Fig. 6.3 A multispectral image of a turmeric powder sample and (b) its spectral signature

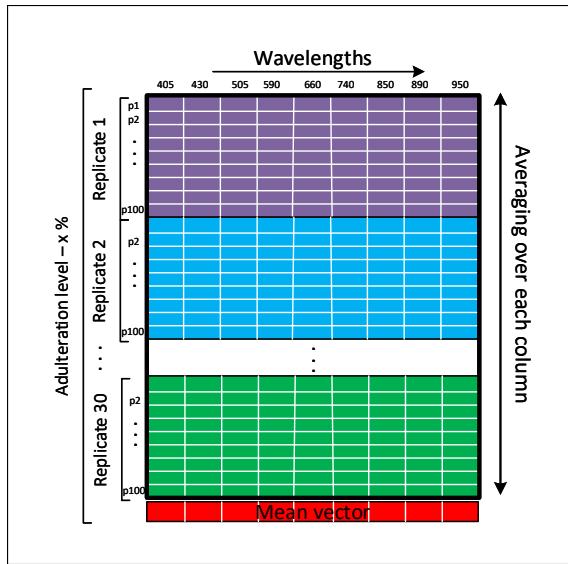


Fig. 6.4 A Two-dimensional data matrix corresponding to an adulteration level. There are altogether nine data matrixes (for the calibration data set) for the adulteration levels of 0%, 5%, 10%, 15%, 20%, 25%, 30%, 35% and 40%

6.3.3 Principal Component Analysis (PCA)

PCA is a statistical procedure, that converts a set of correlated variables (in our case set of wavelengths) into a set of linearly uncorrelated variables called Principal Components (PCs). It is often used for feature extraction and dimensionality reduction problems. Most of the research work carried out in the past used PCA scores to model the system under different external parameters. In our analysis, PCA was used to transform a set of correlated variables (nine wavelengths) into a set of linearly uncorrelated variables without reducing the dimension of the feature space. At the same time, the transformation matrix/feature matrix (P) was created from calibration samples (the set of 270 samples) and the same feature vector (P) matrix was used to transform the validation data-set into principal component space.

In the first step, a matrix $X(m \times n)$ was formulated, where, $X = [X_0, X_5, X_{10}, X_{15}, X_{20}, X_{25}, X_{30}, X_{35}, X_{40}]^T$, where $m = 3000 \times 30$ (No. of replicates) $\times 9$ (No. of adulteration levels) $= 27,000$ and $n = 9$ (No. of wavelengths). Secondly, PCA was performed on the data matrix X using the equations [62, 63] given by,

$$\mu = \frac{1}{m} \sum_{\alpha=1}^{\alpha=m} (X^\alpha)^T \quad (6.7)$$

$$C = E\{[X^\alpha - \mu][X^\alpha - \mu]^T\} \quad (6.8)$$

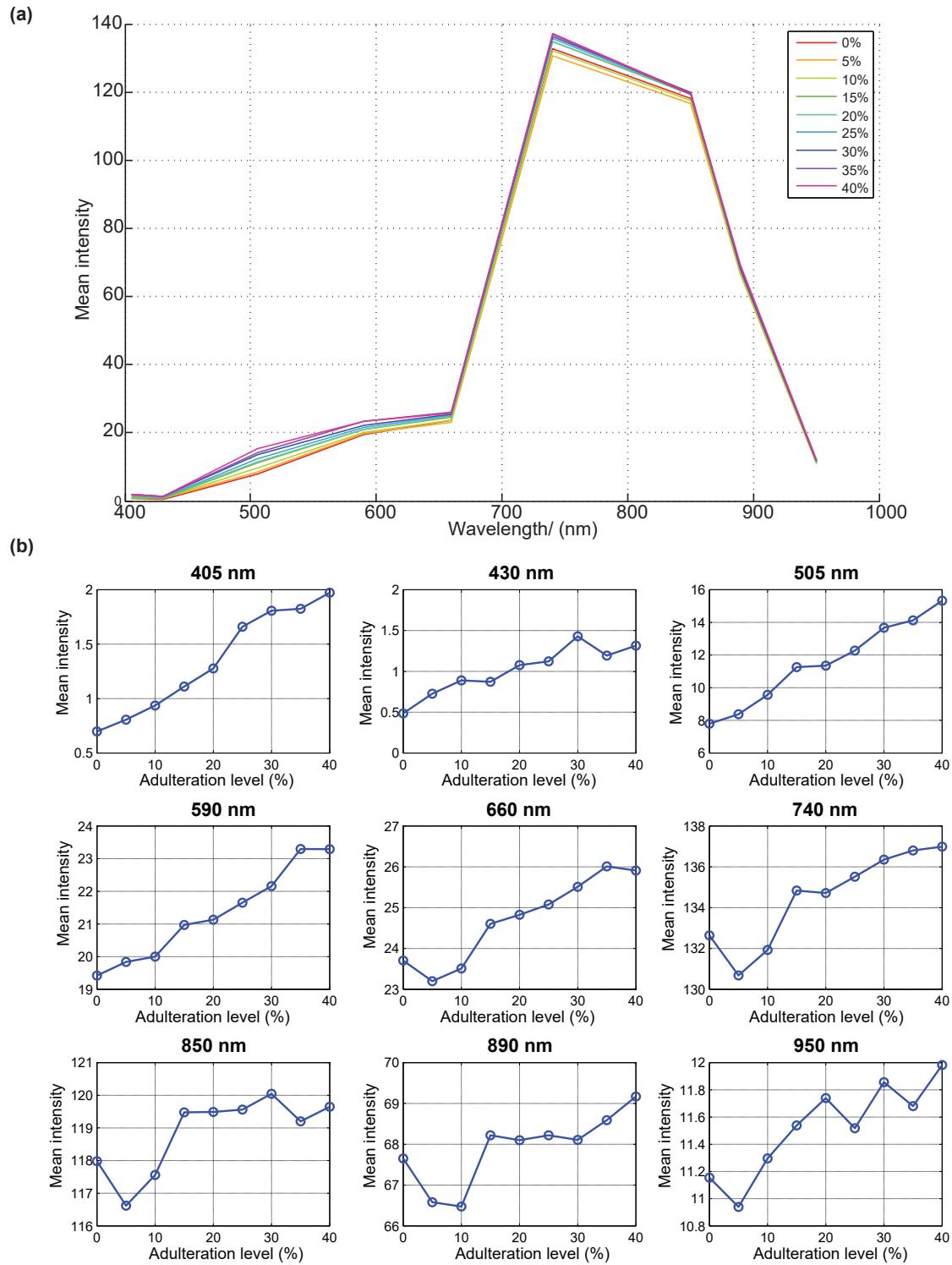


Fig. 6.5 (a) Variation of mean spectral signature curves at different adulteration levels of turmeric powder (b) Variation of mean intensity with adulteration level for different wavelengths

$$C = \frac{1}{m} \sum_{i=1}^{i=m} (X_\alpha - \mu)(X_\alpha - \mu)^T = \bar{X}^T \bar{X} \quad (6.9)$$

$$Cv_k = \lambda_k v_k; \text{ for } k = 1, \dots, n \quad (6.10)$$

$$P = [v_1, v_2, \dots, v_n] \quad (6.11)$$

$$Y = P^T \bar{X} \quad (6.12)$$

Where, $X^\alpha (1 \times n)$ is the vector representation of α^{th} pixel (where $\alpha = 1, 2, \dots, m$), $\mu (n \times 1)$ is the vector representation of mean spectral intensity of all the pixels, $C (n \times n)$ is the covariance matrix, E is the expectation operator, \bar{X} is the mean removed data matrix of X , $v (n \times 1)$ is the eigenvector of the covariance matrix C , λ (scalar) is the eigenvalue of the covariance matrix C , P is the transformation matrix and $Y (m \times n)$ is the transformed/projected data in the new n dimensional space. Finally, the transformed data matrix (Y) was used for the future analysis. It is important to note that, the same transformation matrix/ feature matrix (P) was used to project the validation data-set in to new principal component space (Y_{val}).

6.3.4 Bhattacharyya distance and Jeffries-Matusita (JM) separability measure

Jeffries-Matusita (JM) distance measures the difference between two multivariate Gaussian distributions. It was observed that the data-sets correspond to each adulteration level ($[Y_0, Y_5, Y_{10}, Y_{15}, Y_{20}, Y_{25}, Y_{30}, Y_{35}, Y_{40}]^T = Y = P^T X$) in the principal component space ($PC_1, PC_2, PC_3, \dots, PC_n$, where $n = 9$) forms approximately Gaussian distributions along each principal component axis. Therefore, we can assume that the probability distributions corresponding to each adulteration level are multivariate Gaussian distributed. With this assumption, the dissimilarity between two probability distribution functions (Lets call f_{Y_i} and f_{Y_j} ; where $i = 0$ (pure turmeric) and $j = 0, 5, 10, \dots, 35, 40$ (adulteration levels)) can be measured using the Jeffries-Matusita (JM) distance $J(f_{X_i}, f_{X_j})$ [64, 65] given by,

$$J(f_{Y_i}, f_{Y_j}) = \sqrt{2 \left(1 - e^{-B(f_{Y_i}, f_{Y_j})} \right)} \quad (6.13)$$

$$B(f_{Y_i}, f_{Y_j}) = \frac{1}{8} (\mu_i - \mu_j)^T \left[\frac{C_i + C_j}{2} \right]^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left(\frac{\left| \frac{C_i + C_j}{2} \right|}{\sqrt{|C_i||C_j|}} \right) \quad (6.14)$$

Where $J(f_{Y_i}, f_{Y_j})$ is the Jeffries–Matusita (JM) separability measure between the two multi-variate Gaussian distribution functions f_{Y_i} and f_{Y_j} , $B(f_{Y_i}, f_{Y_j})$ is the Bhattacharyya distance

between the two Gaussian distribution functions f_{Y_i} and f_{Y_j} , $\mu_i(n \times 1)$, $\mu_j(n \times 1)$ are the mean vectors and $C_i(n \times n)$, $C_j(n \times n)$ are the covariance matrixes. The range of $J(f_{Y_i}, f_{Y_j})$ from 0 to 2. When a turmeric powder sample contains adulterants, it's multivariate distribution function differs from the authentic turmeric sample. The degree of dissimilarity between the two distributions depend on the adulteration level. Therefore, from a mathematical point of view, when more and more adulterants are mixed with turmeric powder, its multivariate distributions differ more and more from the pure turmeric and as a result, probabilistic distance between two multivariate distribution functions will increase significantly. Further, according to the value of $J(f_{Y_i}, f_{Y_j})$, we were able to categorized the quality of turmeric powder samples into three regions,

$$J(f_{Y_i}, f_{Y_j}) = \begin{cases} GradeC & \text{if } 2.0 \geq J \geq 1.4 \\ GradeB & \text{if } 1.4 > J \geq 0.6 \\ GradeA & \text{if } 0.6 > J \geq 0 \end{cases} \quad (6.15)$$

Here, if a turmeric powder sample falls into Grade A, it means its properties are more identical to the authentic turmeric powder (adulteration level $< 5\%$), if a turmeric powder sample falls into Grade B that means it contains $< 15\%$ of adulterants and if a turmeric powder sample falls into Grade C it contains $> 15\%$ of adulterants.

6.3.5 Maximum Likelihood Estimation (MLE)

In our research, MLE has been used to cross-validate the results obtained from Jeffries-Matusita (JM) separability measure and Bhattacharyya distance. The Maximum Likelihood Estimation is a supervised classification method derived using the Bayes theorem [64]. According to the previous notations, X_0 denotes the data matrix corresponding to the pure turmeric powder samples in which there are data points selected from thirty replicates. Let's denote, k^{th} ($1 \leq k \leq m$) data point selected from l^{th} ($1 \leq l \leq 30$) replicate of the data matrix - X_0 as $x_{0,l}^{(k)}$. Assuming the probability density functions of all replicates are gaussian distributed, the probability density functions of each replicate can be expressed as,

$$f_{X_0}(X_{0,l}) = \frac{1}{\sqrt{2\pi|C_l|}} e^{-\frac{1}{2}((\omega - \mu_l)^T C_l^{-1} (\omega - \mu_l))} \quad (6.16)$$

where,

$$\mu_l = \frac{1}{m} \sum_{k=1}^{k=3000} (x_{0,l}^k - \mu_l)^T (x_{0,l}^k - \mu_l) \quad (6.17)$$

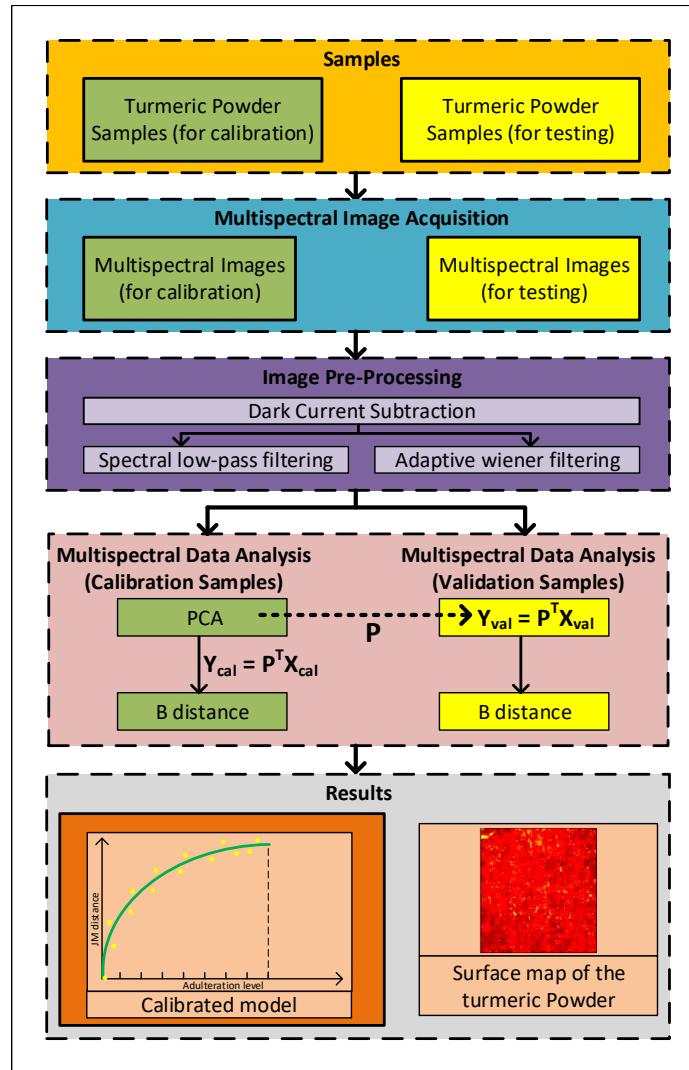


Fig. 6.6 Complete algorithm as a schematic

Therefore, the likelihood function can be written as,

$$g(\omega) = \frac{1}{30} \sum_{l=1}^{l=30} \frac{1}{\sqrt{2\pi |C_l|}} e^{-\frac{1}{2}((\omega - \mu_i)^T C_l^{-1} (\omega - \mu_i))} \quad (6.18)$$

Then, the normalized log likelihood function value can be plotted as a spatial map. The summary of the proposed algorithm is shown in Figure 6.6.

Table 6.2 The performance of curve fitting results under the Wiener filtering (window size: 3 pixels \times 3 pixels) and the median filtering (window size: 3 pixels \times 3 pixels)

Curve fitting result: $B = p_1p^2 + p_2p + p_3$ and p_3 fixed at zero.					
	Squared Error of prediction (SSE)	Coefficient of determination (R^2)	Root Mean Square Error (RMSE)	Coefficients of the polynomial	
				P1	P2
The Wiener filter: [3 pixels \times 3 pixels]	0.0500	0.9911	0.08451	0.001007	0.02077
The median filter: [3 pixels \times 3 pixels]	0.1772	0.9866	0.15910	0.001876	0.01801

6.4 Results and Discussion

6.4.1 Selection of the optimal filter and filter parameters

To remove the random noise present in a multispectral image, two types of filtering methods were tested separately at the image pre-processing step: the median filtering and the Wiener filtering. The best filter type and the window size (n pixels \times n pixels) were determined such that it maximizes the coefficient of determination (R^2) for the curve fitting results. If the size of the window is too large, the filter will blur the image and sometimes it will result in disappearing some of the useful information present in the multispectral image. On the other hand, if the window size is too small, random noise will still be present in the image. Therefore, a reasonable window size (3 pixels \times 3 pixels) was selected. For the 3 pixels \times 3 pixels window size, the median filtering resulted in an R^2 value of 0.9866 and the Wiener filtering resulted in an R^2 value of 0.9911 as shown in Figure 6.7. The performance of the curve fitting results under these two different filtering methods are summarized in Table 6.2. Finally, the adaptive Wiener filtering with the window size of 3 pixels \times 3 pixels was performed at the image pre-processing step following the dark current subtraction.

6.4.2 Prediction of Adulteration Level

The curve fitting results for the Bhattacharyya Distance (B) and the JM distance with the adulteration level are shown in Figure 6.8 - (a), (b) respectively. The variation of the Bhattacharya distance (B) with the adulteration level was modeled using a second-order polynomial ($B = p_1P^2 + p_2P + P_3$). The coefficients of the polynomial were calculated as $P_1 = 0.001007$, $P_2 = 0.02077$ and $P_3 = 0$ with an R^2 value of 0.9911. Then the variation of the JM distance with the adulteration level was modeled using the equation given in Equation 6.13. Next, this model was validated using a set of validation samples. Sixty samples with adulteration levels of $p(i)$ given by the equation $p(i) = \frac{0.1i}{9+0.1i} \times 100\%$ (where $i = 0, 1, 2, \dots, 59$) were used as the validation set. Good prediction accuracy was obtained for the validation set with an R^2 value of 0.9644 as shown in Figure 6.9.

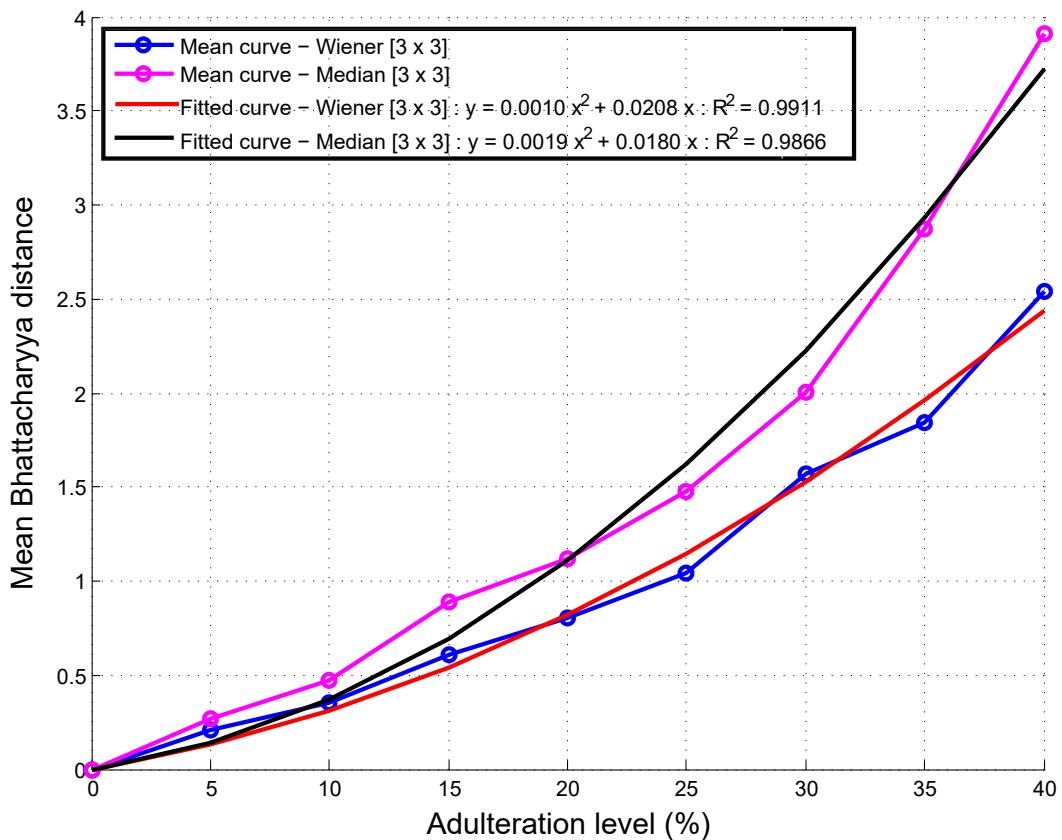


Fig. 6.7 The Curve fitted results (variation of mean Bhattacharyya Distance with adulteration level) for the Wiener filtering with window size of 3 pixels \times 3 pixels (blue and pink lines) and for the median filtering with window size of 3 pixels \times 3 pixels (red and black lines)

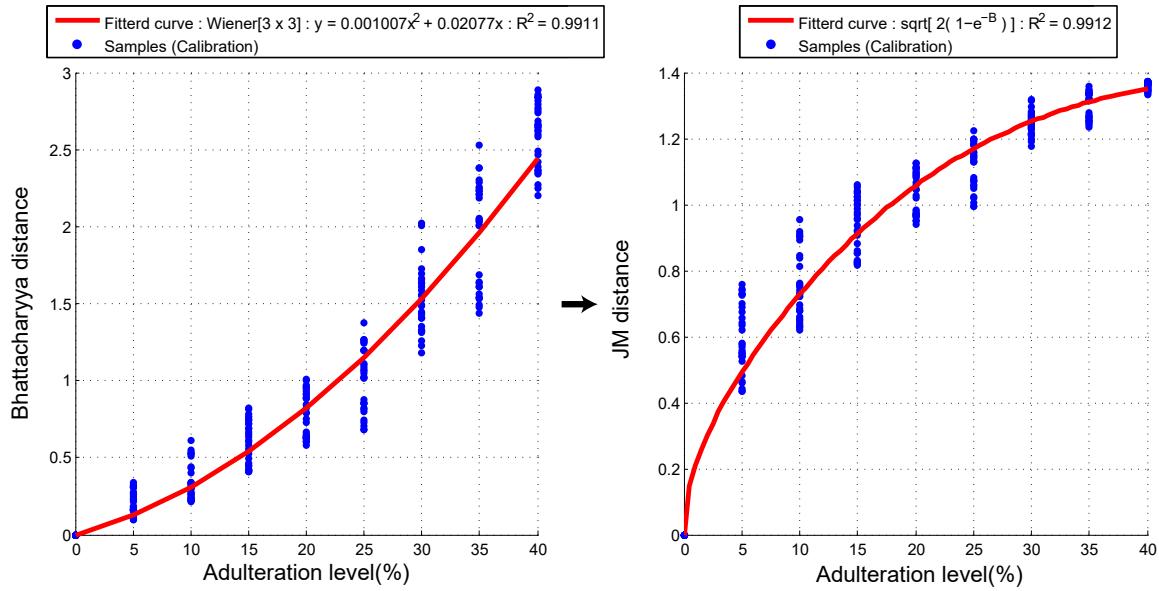


Fig. 6.8 The Curve fitted results for training data (a) variation of the Bhattacharya Distance with adulteration level (b) variation of the JM Distance with adulteration level

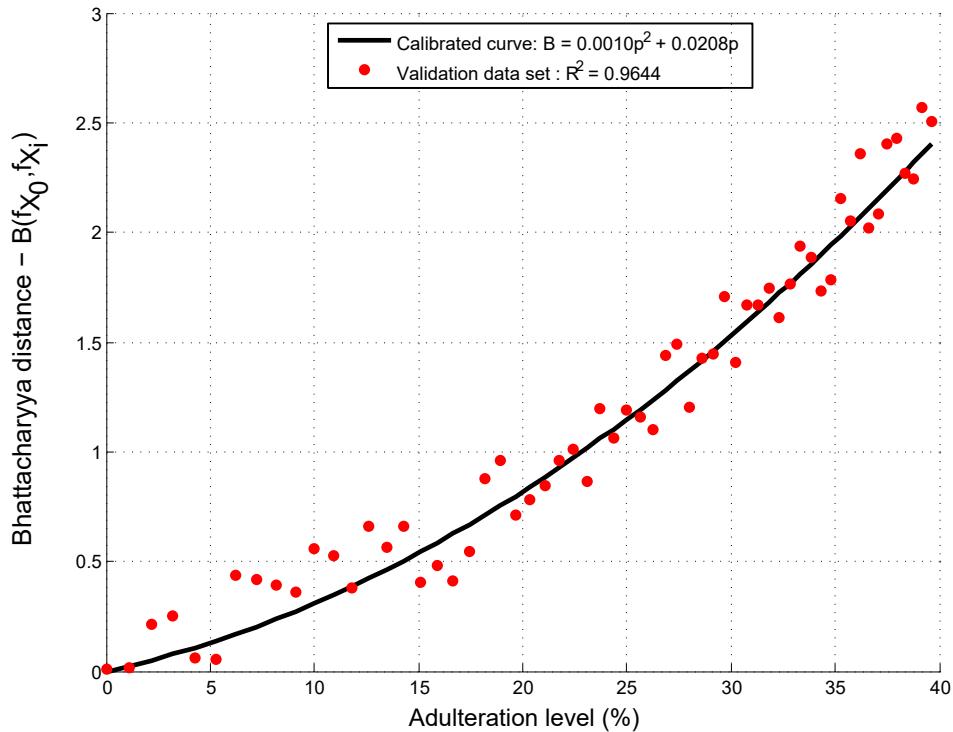


Fig. 6.9 The variation of Bhattacharyya Distance with adulteration level for validation data on the calibrated curve ($B = 0.0010p^2 + 0.0208p$). The correlation coefficient between calibrated curve and the validation data is 0.9644

Table 6.3 Details of estimated adulteration levels of commercial samples

Commercial product	Bhattacharyya distance - B					Mean Bhattacharyya distance - B	Variance	Standard Deviation	Predicted adulteration level (P)
	Replicate 1	Replicate 2	Replicate 3	Replicate 4	Replicate 5				
CS1	0.6300	0.6407	0.6743	0.6662	0.6160	0.6454	0.00060	0.0244	17.0%
CS2	0.5723	0.5503	0.5734	0.5264	0.5264	0.5497	0.00054	0.0232	15.2%
CS3	0.0765	0.1236	0.0786	0.0742	0.0797	0.0865	0.00043	0.0209	3.6%
CS4	0.0682	0.0623	0.0643	0.0720	0.0602	0.0654	0.00002	0.0047	2.8%
CS5	0.0797	0.0797	0.0797	0.0797	0.0797	0.0797	0.00000	0.0001	3.3%
CS6	0.0969	0.0991	0.0872	0.1027	0.1027	0.0977	0.00004	0.0064	3.9%
CS7	0.1127	0.1248	0.1190	0.1163	0.1163	0.1178	0.00002	0.0045	4.6%
CS8	0.1103	0.1117	0.1036	0.1036	0.1103	0.1079	0.00001	0.0040	4.3%
CS9	0.1167	0.1177	0.1164	0.1099	0.1099	0.1141	0.00001	0.0039	4.5%
CS10	0.0595	0.0673	0.0671	0.0655	0.0655	0.0650	0.00001	0.0032	2.8%

To further validate the results, the adulteration levels of the four commercially available products were determined using our calibrated model and the results were verified from a laboratory test. These four commercially available products were named as *CS₁*, *CS₂*, *CS₃*, and *CS₄* for easy recognition. Then five replicates were taken from each product and the Bhattacharyya distances were calculated for each replicate. After that, the mean Bhattacharyya distance was used to estimate the adulteration level of each of the samples as shown in Table 3. According to our calculations, Product *CS₁* contained adulterants with approximately 17% (w/w) and Products *CS₂*, *CS₃* and *CS₄* contained adulterants of about 5% (w/w). To validate the result, a small quantity from each product was analyzed using light microscope. The micrographs of turmeric obtained under the magnification 10 × 10 are shown in Figure 12.

As shown in Figure 6.11 - (c), the sample *CS₁* contained large particles compared to authentic turmeric powder particles. Further, the large particles were identified as rice scratch by carefully observing the microscopic views and performing an iodine test. Even though our calibrated model uses tartrazine and wheat flour as the adulterants, the adulteration level estimated from the model is reasonably accurate. When the adulterants were mixed with pure turmeric, its probability distribution deviates from the probability distribution of the authentic turmeric. Hence the probabilistic distance between the two distributions will be increased. Because of this reason, any turmeric powder sample mixed with adulterants can easily be identified using the proposed method. Figure 6.11- (d), (e) and (f) show the microscopic views of samples *CS₂*, *CS₃*, and *CS₄* respectively. By carefully observing the above figures, we can identify that these three products do not contain any adulterants but contain some impurities like adventitious roots, remains of the leaves, vascular bundles and old parts of the stems with a little amount as estimated from our model. Finally, the adulteration levels predicted from our calibrated model were accurately matched with the results obtained using a microscopic test and the iodine analysis.

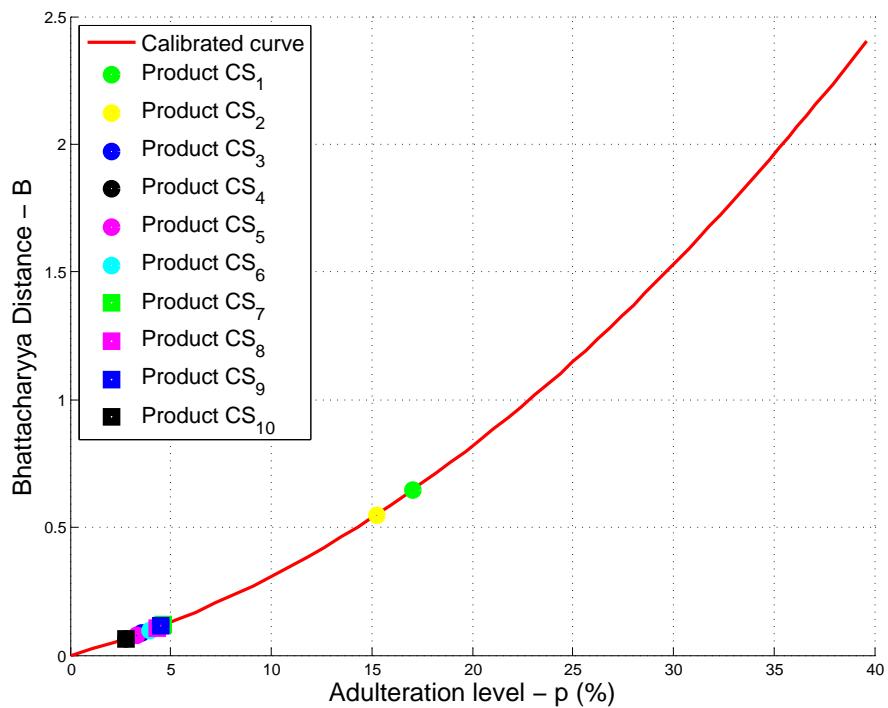


Fig. 6.10 The Bhattacharyya distance of commercially available samples CS_1 , CS_2 , CS_3 , CS_4 , CS_5 , CS_6 , CS_7 , CS_8 , CS_9 and CS_{10} on the calibration curve

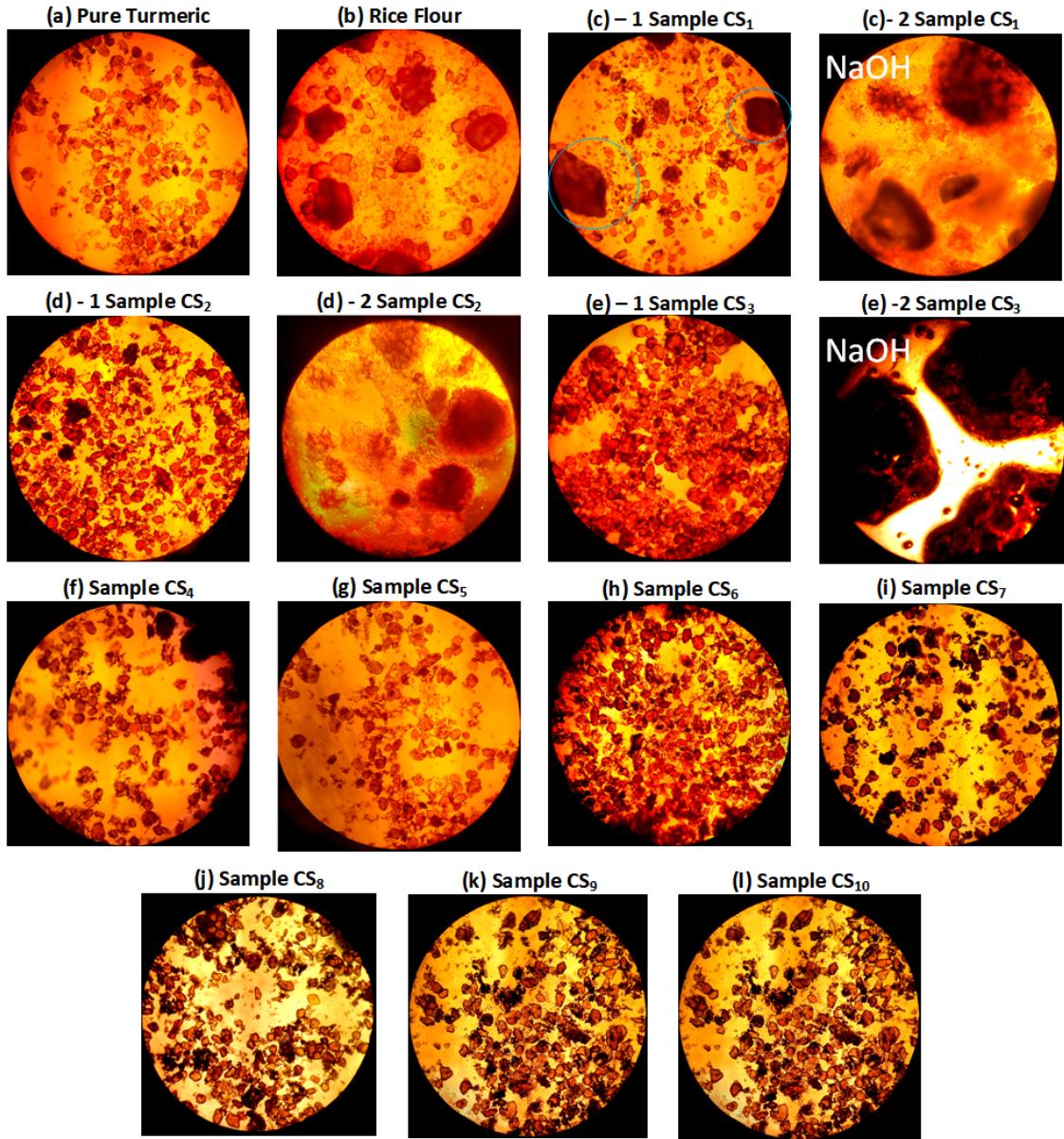


Fig. 6.11 Microscopic views of (a) authentic turmeric, (b) rice flour, and (c) – (i) commercially available turmeric samples CS_1 , CS_2 , CS_3 , CS_4 , CS_5 , CS_6 , CS_7 , CS_8 , CS_9 and CS_{10} respectively. Magnification: 10×10 .

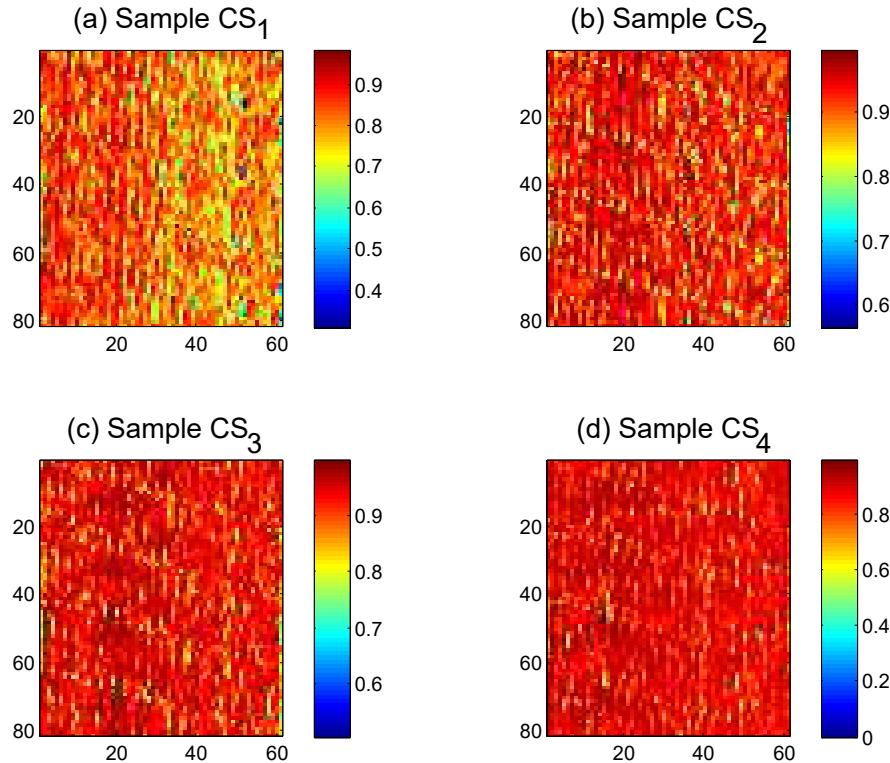


Fig. 6.12 Variability in normalized log MLE score over the surface of turmeric powder samples (a) Sample CS_1 (b) Sample CS_2 (c) Sample CS_3 (d) Sample CS_4

The spatial distribution (map) of the normalized average log-likelihood scores (from Eq. 19) were plotted for each commercially available sample as shown in Figure 13. From the color charts, we can predict whether each pixel belongs more likely to pure turmeric or not. If a pixel has a color closer to dark red, it means that there is a higher probability that the pixel belongs to authentic turmeric. In contrast, if the pixel is more likely to be an adulterant or an impurity, then the color of the pixel deviates from dark red to a color closer to yellow, green and blue. As shown in Figure 13 – (a), the yellow color pixels are more dominant because, as discussed previously, the product CS_1 contained approximately 17% of adulterants. Similarly, the surface map of the products CS_2 , CS_3 , and CS_4 are closer to dark red, meaning that the pixels are more likely to be authentic turmeric as expected.

6.5 Conclusion

Results of this study confirm the possibility of applying the Bhattacharya distance followed by PCA to determine the adulteration level of turmeric powder using multispectral imaging. A multispectral imaging system was developed with nine spectral bands selected from UV to NIR region of the electromagnetic spectrum. A mixture of tartrazine and wheat flour was used as the adulterant; one of the most common adulterants found in turmeric. The calibrated model was developed with an R^2 value of 0.9911. Then the model was validated using a set of validation samples with $R^2 = 0.9644$ and SSE = 1.0569. Moreover, the model was used to estimate the adulteration content of four commercially available products and the results were verified by microscopic analysis and an iodine test. Even though this study was limited to tartrazine and wheat flour as the adulterants, depending on the requirement the algorithm can be modified such that it enables distinguishing between the adulterants. This can be achievable using a reasonable set of calibration samples for each adulterant.

References

- [1] Electromagnetic spectrum. [Online]. Available: <https://sites.google.com/site/chempendix/em-spectrum>
- [2] J. Klein, J. Brauers, and T. Aach, “Spatio-spectral modeling and compensation of transversal chromatic aberrations in multispectral imaging,” vol. 55, pp. 60 502–1, 11 2011.
- [3] R. Ennis, F. Schiller, M. Toscani, and K. R. Gegenfurtner, “Hyperspectral database of fruits and vegetables,” *J. Opt. Soc. Am. A*, vol. 35, no. 4, pp. B256–B266, Apr 2018. [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-35-4-B256>
- [4] G. ElMasry and D.-W. Sun, “Chapter 1 - principles of hyperspectral imaging technology,” in *Hyperspectral Imaging for Food Quality Analysis and Control*, D.-W. Sun, Ed. San Diego: Academic Press, 2010, pp. 3 – 43. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/B9780123747532100012>
- [5] C. U. Computer Vision Laboratory, Department of Computer Science. Cave | multispectral image database: Real and fake. [Online]. Available: http://www.cs.columbia.edu/CAVE/databases/multispectral/real_and_fake/
- [6] W. G. C. Bandara, G. W. K. Prabhath, D. W. S. C. B. Dissanayake, V. R. Herath, G. M. R. I. Godaliyadda, M. P. Ekanayake, S. Vithana, D. Demini, and T. Madhujith, “A multispectral imaging system to assess meat quality,” in *2018 IEEE Region 10 Humanitarian Technology Conference (R10-HTC) (R10 HTC'18)*, Colombo, Sri Lanka, Dec. 2018.
- [7] P. . T. CORP. (2016) Multispectral imaging and it’s applications. [Online]. Available: <https://www.phase1vision.com/blog/multispectral-imaging-and-its-applications>
- [8] H. Yao and D. Lewis, “Chapter 2 - spectral preprocessing and calibration techniques,” in *Hyperspectral Imaging for Food Quality Analysis and Control*, D.-W. Sun, Ed. San Diego: Academic Press, 2010, pp. 45 – 78. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/B9780123747532100024>
- [9] D. G. Manolakis, R. B. Lockwood, and T. W. Cooley, *Linear Spectral Transformations*. Cambridge University Press, 216, pp. 406–422. [Online]. Available: <https://www.cambridge.org/core/books/hyperspectral-imaging-remote-sensing/E19050621F91BFCAF904CD1A8CB35F29>

- [10] T. Raiko, A. Ilin, and J. Karhunen, “Principal component analysis for large scale problems with lots of missing values,” in *Machine Learning: ECML 2007*, J. N. Kok, J. Koronacki, R. L. d. Mantaras, S. Matwin, D. Mladenić, and A. Skowron, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 691–698.
- [11] T. Oorloff, A. Abeysekara, S. Vithana, R. Rupasinghe, H. Herath, G. Godaliyadda, and M. Ekanayake, “Spectral-spatial hybrid mechanism for feature detection using spectral correlation,” in *2017 IEEE International Conference on Industrial and Information Systems (ICIIS)*, Dec 2017, pp. 1–6.
- [12] S. Vithana, R. Abeysekara, S. Oorloff, V. Herath, R. Godaliyadda, P. Ekanayake, and A. Rupasinghe, “Comparison of two algorithms for land cover mapping based on hyperspectral imagery,” *ICTer Journal*, vol. 11, no. 1, 2018.
- [13] J. Zabalza, J. Ren, M. Yang, Y. Zhang, J. Wang, S. Marshall, and J. Han, “Novel folded-pca for improved feature extraction and data reduction with hyperspectral imaging and sar in remote sensing,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, pp. 112 – 122, 2014.
- [14] M. Imani and H. Ghassemian, “Feature space discriminant analysis for hyperspectral data feature reduction,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 102, pp. 1 – 13, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0924271615000179>
- [15] A. Y. Ng, M. I. Jordan, and Y. Weiss, “On spectral clustering: Analysis and an algorithm,” in *Advances in Neural Information Processing Systems 14*, T. G. Dietterich, S. Becker, and Z. Ghahramani, Eds. MIT Press, 2002, pp. 849–856. [Online]. Available: <http://papers.nips.cc/paper/2092-on-spectral-clustering-analysis-and-an-algorithm.pdf>
- [16] R. Shrestha and J. Y. Hardeberg, “Multispectral imaging using led illumination and an rgb camera,” *Color and Imaging Conference*, vol. 2013, no. 1, pp. 8–13, 2013.
- [17] C. Matasaru, “Master erasmus mundus in color in informatics and media technology (cimet) mobile phone camera possibilities for spectral imaging master thesis report,” 2014.
- [18] M. Goel, E. Whitmire, A. Mariakakis, T. S. Saponas, N. Joshi, D. Morris, B. Guenter, M. Gavriliu, G. Borriello, and S. N. Patel, “Hypercam: Hyperspectral imaging for ubiquitous computing applications,” in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp ’15. New York, NY, USA: ACM, 2015, pp. 145–156.
- [19] “Tech guide: Integrating sphere radiometry and photometry - labsphere | internationally recognized photonics company.” [Online]. Available: <https://www.labsphere.com/support/system-product-brochures/tech-guide-a-guide-to-integrating-sphere-radiometry-and-photometry/>
- [20] “Chapter 2 - spectral preprocessing and calibration techniques,” in *Hyperspectral Imaging for Food Quality Analysis and Control*, D.-W. Sun, Ed. San Diego: Academic Press, 2010, p. 61.

- [21] "Determining quality and maturity of pomegranates using multispectral imaging," *Journal of the Saudi Society of Agricultural Sciences*, vol. 16, no. 4, pp. 322 – 331, 2017.
- [22] G. ElMasry and D.-W. Sun, "Chapter 6 - meat quality assessment using a hyperspectral imaging system," in *Hyperspectral Imaging for Food Quality Analysis and Control*, D.-W. Sun, Ed. San Diego: Academic Press, 2010, pp. 179–183. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/B9780123747532100061>
- [23] R. Dutson, "The measurement of ph in muscle and its importance to meat quality thayne," 2003.
- [24] X. Li, J. Babol, A. Wallby, and K. Lundström, "Meat quality, microbiological status and consumer preference of beef gluteus medius aged in a dry ageing bag or vacuum," vol. 95, pp. 229–234, 05 2013.
- [25] Y. Liu, B. G. Lyon, W. R. Windham, C. E. Realini, T. D. Pringle, and S. Duckett, "Prediction of color, texture, and sensory characteristics of beef steaks by visible and near infrared reflectance spectroscopy. a feasibility study," *Meat Science*, vol. 65, no. 3, pp. 1107 – 1115, 2003.
- [26] N. Elsharawy, A. Meawad Ahmad, and H. A Abdelrahman, "Quality assessment of nutritional value and safety of different meat," vol. 03, 01 2018.
- [27] M. Larry and T. P. James, "Bacteriological analytical manual : Chapter 3 :aerobic plate count," Jan 2001. [Online]. Available: <https://www.fda.gov/food/foodscienceresearch/laboratorymethods/ucm063346.htm>
- [28] G. K. Jayaprakasha, L. J. M. Rao, and K. K. Sakariah, "Improved HPLC method for the determination of curcumin, demethoxycurcumin, and bisdemethoxycurcumin," *Journal of Agricultural and Food Chemistry*, vol. 50, no. 13, pp. 3668–3672, 2002.
- [29] A. J. Rubya, K. D. Babub, and K. N. Rajasekharanb, "Cancer letters," *Cancer Letters*, vol. 94, pp. 79–83, 1995.
- [30] B. Joe, M. Vijaykumar, and B. R. Lokesh, "Biological Properties of Curcumin-Cellular and Molecular Mechanisms of Action," *Critical Reviews in Food Science and Nutrition*, vol. 44, no. 2, pp. 97–111, 2004.
- [31] S. Naz, S. Jabeen, S. Ilyas, F. Manzoor, F. Aslam, and A. Ali, "Antibacterial activity of Curcuma longa varieties against different strains of bacteria," *Pakistan Journal of Botany*, vol. 42, no. 1, pp. 455–462, 2010.
- [32] M. Chainani-wu Nita, D.M.D., M.P.H., "Safety and Anti-Inflammatory Activity of Curcumin ;," *The Journal of Alternative and Complementary Medicine*, vol. 9, no. 1, pp. 161–168, 2003.
- [33] *No Title - .*
- [34] S. Daniells, "Adulterated turmeric found in supermarket chainrsquos sample," 2018.
- [35] Pranali Lotlikar Chindarkar, "Sessions court upholds action against Mumbai-based firm for food adulteration," 2018. [Online]. Available: <http://dnai.in/fQKC>

- [36] S. R. Bhatt, "Impact analysis of food adulterants on health in some selected urban area of Varanasi," Ph.D. dissertation, 2015.
- [37] S. Dhakal, K. Chao, W. Schmidt, J. Qin, M. Kim, and D. Chan, "Evaluation of Turmeric Powder Adulterated with Metanil Yellow Using FT-Raman and FT-IR Spectroscopy," *Foods*, vol. 5, no. 2, p. 36, 2016. [Online]. Available: <http://www.mdpi.com/2304-8158/5/2/36>
- [38] R. Shah, "Identification and Estimation of Non-Permitted Food Colours (Metanil Yellow and Aniline Dyes) in Turmeric Powder By Rapid Color Test and Thin Layer Chromatography," *World Journal of Pharmacy and Pharmaceutical Sciences*, vol. 6, no. 8, pp. 2034–2045, 2017. [Online]. Available: http://wjpps.com/wjpps{_}controller/abstract{_}id/7605
- [39] R. W. L.W.D. van Raamsdonk, "Microscopy for fraud detection," Tech. Rep., 2015. [Online]. Available: https://www.wur.nl/upload{_}mm/6/2/1/e8e93964-f0a9-4b9a-9aab-1f2c6ca1261f{_}5.VanRaamsdonk{_}Spiced-Workshop{_}2015.pdf
- [40] A. R. Sen, P. Sen Gupta, and N. Chose Dastidar, "Detection of Curcuma zedoaria and Curcuma aromatica in Curcuma longa (Turmeric) by thin-layer chromatography," *The Analyst*, vol. 99, no. 1176, pp. 153–155, 1974.
- [41] V. Ashok, N. Agrawal, A. Durbanshi, J. Esteve-Romero, and D. Bose, "A novel micellar chromatographic procedure for the determination of metanil yellow in foodstuffs," *Analytical Methods*, vol. 7, no. 21, pp. 9324–9330, 2015.
- [42] M. R. Fuh and K. J. Chia, "Determination of sulphonated azo dyes in food by ion-pair liquid chromatography with photodiode array and electrospray mass spectrometry detection," *Talanta*, vol. 56, no. 4, pp. 663–671, 2002.
- [43] A. I. Ropodi, E. Z. Panagou, and G. J. E. Nychas, "Multispectral imaging (MSI): A promising method for the detection of minced beef adulteration with horsemeat," *Food Control*, vol. 73, pp. 57–63, 2017.
- [44] D. Bertelli, M. Plessi, A. G. Sabatini, M. Lolli, and F. Grillenzoni, "Classification of Italian honeys by mid-infrared diffuse reflectance spectroscopy (DRIFTS)," *Food Chemistry*, vol. 101, no. 4, pp. 1565–1570, 2007.
- [45] R. Khodabakhshian, B. Emadi, M. Khojastehpour, and M. R. Golzarian, "Determining quality and maturity of pomegranates using multispectral imaging," *Journal of the Saudi Society of Agricultural Sciences*, vol. 16, no. 4, pp. 322–331, 2017. [Online]. Available: <https://doi.org/10.1016/j.jssas.2015.10.004>
- [46] W.-h. Su and D.-w. Sun, "RAPID AND NONDUSTRUCTIVE DETECTION OF ADULTERANTS OF ORGANIC SPELT FLOUR USING."
- [47] M. Goel, E. Whitmire, A. Mariakakis, T. S. Saponas, N. Joshi, D. Morris, B. Guenter, M. Gavriliu, G. Borriello, and S. N. Patel, "Hyper Cam : Hyperspectral Imaging for Ubiquitous Computing Applications."

- [48] J. Aguzzi, D. P. Ariana, J. Blasco, K. Chao, C. Costa, G. ElMasry, J. Gómez-Sanchis, A. A. Gowen, D. S. Jayas, L. Jiang, D. Lewis, L. Liu, R. Lu, P. Menesatti, E. Moltó, M. O. Ngadi, C. P. O'Donnell, J. Paliwal, G. Polder, J. Qin, C. B. Singh, J. Sugiyama, D.-W. Sun, M. Taghizadeh, Y. Tao, M. Tsuta, G. van der Heijden, N. Wang, H. Yao, and B. Zhu, "Contributors," in *Hyperspectral Imaging for Food Quality Analysis and Control*, D.-W. Sun, Ed. San Diego: Academic Press, 2010, pp. ix – xi. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/B978012374753210019X>
- [49] C. Bandara, B. Dissanayake, K. Prabhath, P. Vithana, R. Herath, I. Godaliyadda, and B. Ekanayake, "A Multispectral Imaging System to Assess Meat Quality," 2018.
- [50] S. Baek, E. Choi, Y. Baek, and C. Lee, "Detection of counterfeit banknotes using multispectral images," *Digital Signal Processing: A Review Journal*, vol. 78, pp. 294–304, 2018. [Online]. Available: <https://doi.org/10.1016/j.dsp.2018.03.015>
- [51] M. M. Chaudhry, M. L. Amodio, F. Babellahi, M. L. de Chiara, J. M. Amigo Rubio, and G. Colelli, "Hyperspectral imaging and multivariate accelerated shelf life testing (MASLT) approach for determining shelf life of rocket leaves," *Journal of Food Engineering*, vol. 238, no. June, pp. 122–133, 2018. [Online]. Available: <https://doi.org/10.1016/j.jfoodeng.2018.06.017>
- [52] K. Sendin, M. Manley, and P. J. Williams, "Classification of white maize defects with multispectral imaging," *Food Chemistry*, vol. 243, no. June 2017, pp. 311–318, 2018. [Online]. Available: <http://dx.doi.org/10.1016/j.foodchem.2017.09.133>
- [53] C. Tang, H. He, E. Li, and H. Li, "Multispectral imaging for predicting sugar content of 'Fuji' apples," *Optics and Laser Technology*, vol. 106, pp. 280–285, 2018. [Online]. Available: <https://doi.org/10.1016/j.optlastec.2018.04.017>
- [54] M. Goel, S. N. Patel, E. Whitmire, A. Mariakakis, T. S. Saponas, N. Joshi, D. Morris, B. Guenter, M. Gavriliu, and G. Borriello, "HyperCam," *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*, pp. 145–156, 2015. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=2750858.2804282>
- [55] K. Hoshino, F. Nielsen, and T. Nishimura, "Noise Reduction in CMOS Image Sensors for High Quality Imaging : The Autocorrelation Function Filter on Burst Image Sequences ACF for image processing," *ICGST-GVIP Journal*, vol. 7, no. 3, pp. 3–10, 2007.
- [56] Y. Reibel, M. Jung, M. Bouhifd, B. Cunin, and C. Draman, "CCD or CMOS camera noise characterisation," *The European Physical Journal Applied Physics*, vol. 21, pp. 75–80, 2003.
- [57] Y. Yoo, S. Lee, W. Choe, and C.-y. Kim, "Processor in Digital Cameras and Camera Phones," *Analysis*, vol. 6502, pp. 1–10, 2007.
- [58] Timothy York, "Fundamentals of Image Sensor Performance," 2011. [Online]. Available: <https://www.cse.wustl.edu/{~}jain/cse567-11/ftp/imgsens/index.html>

- [59] W. C. Porter, B. Kopp, J. C. Dunlap, R. Widenhorn, and E. Bodegom, “Dark current measurements in a CMOS imager,” p. 68160C, 2008. [Online]. Available: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.769079>
- [60] “About the Editor,” in *Hyperspectral Imaging for Food Quality Analysis and Control*, D.-W. Sun, Ed. San Diego: Academic Press, 2010, pp. vii – viii. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/B9780123747532100188>
- [61] J. A. E. S. Lim, *AND RADER Number Theory in Digital Signal Processing Seismic Applications of Homomorphic Signal Processing PROCESSING*.
- [62] D. G. Manolakis, R. B. Lockwood, and T. W. Cooley, *Hyperspectral Imaging Remote Sensing: Physics, Sensors, and Algorithms*. Cambridge University Press, 2016.
- [63] J. Zabalza, J. Ren, M. Yang, Y. Zhang, J. Wang, S. Marshall, and J. Han, “Novel Folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, pp. 112–122, 2014. [Online]. Available: <http://dx.doi.org/10.1016/j.isprsjprs.2014.04.006>
- [64] a. Ahmad and S. Quegan, “Analysis of maximum likelihood classification on multispectral data,” *Applied Mathematical Sciences*, vol. 6, no. 129, pp. 6425–6436, 2012.
- [65] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, 3rd ed., D. E. Ricken and W. Gessner, Eds. Berlin, Heidelberg: Springer-Verlag, 1999.

Appendix A

LED Driver Circuit - Schematic and Layout

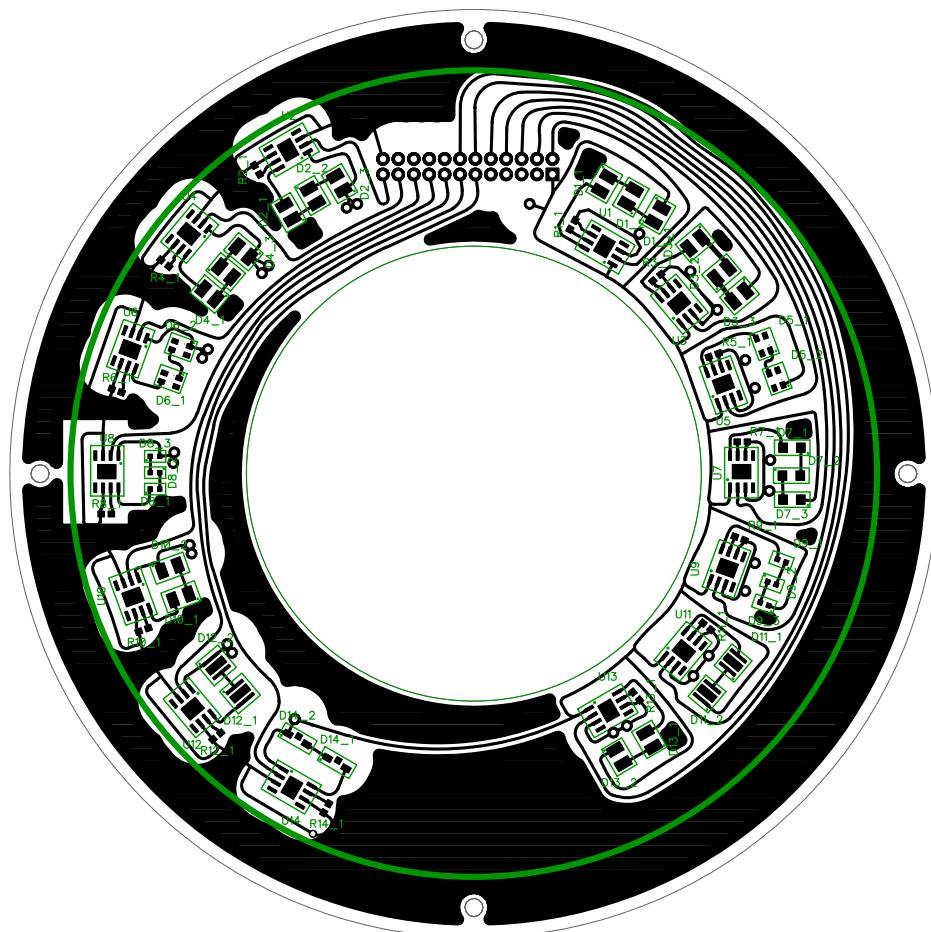


Fig. A.1 Layout of the LED driver circuit

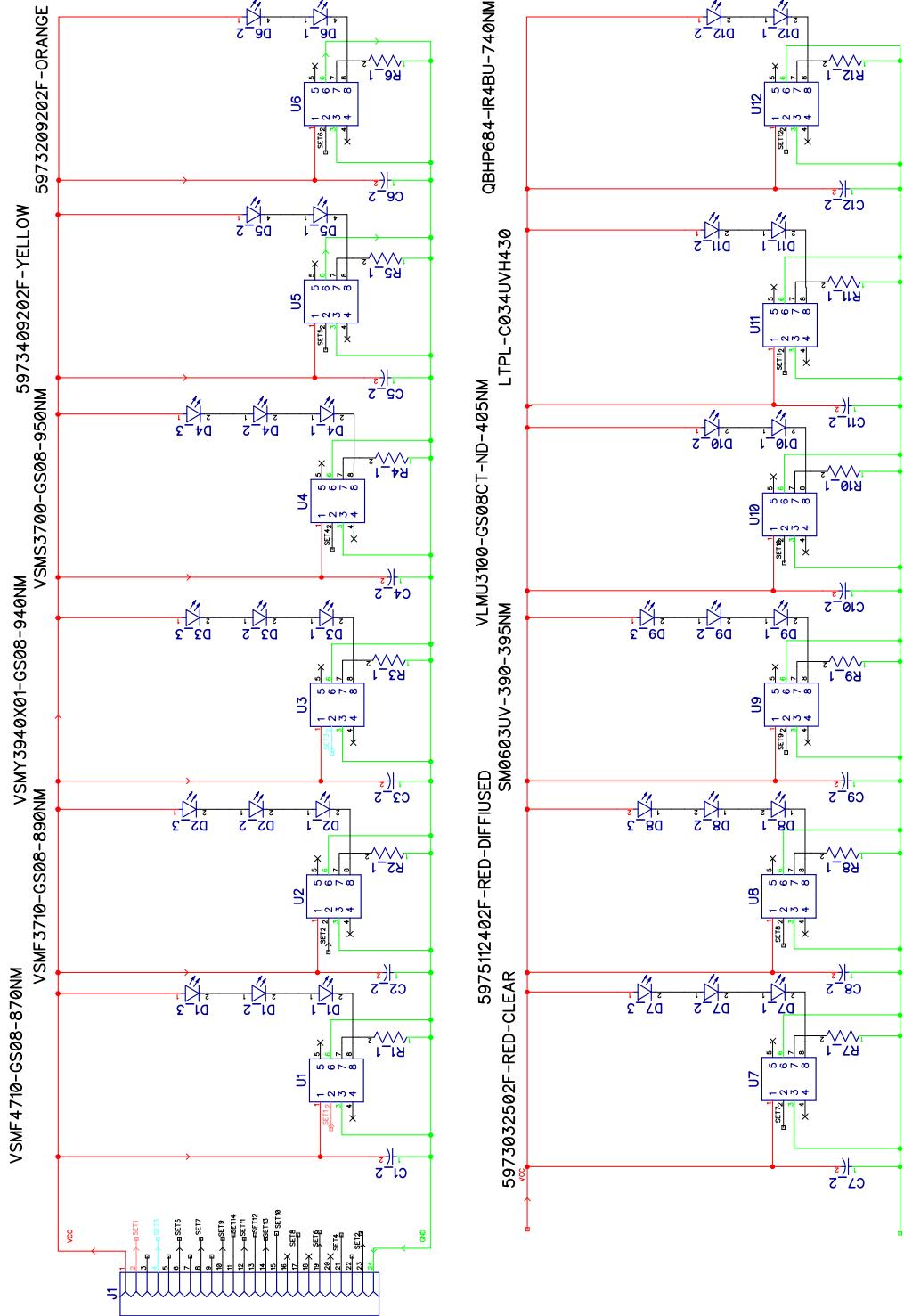


Fig. A.2 Schematic diagram of the LED driver circuit

Appendix B

Discovery Board Support Circuit - Schematic and Layout

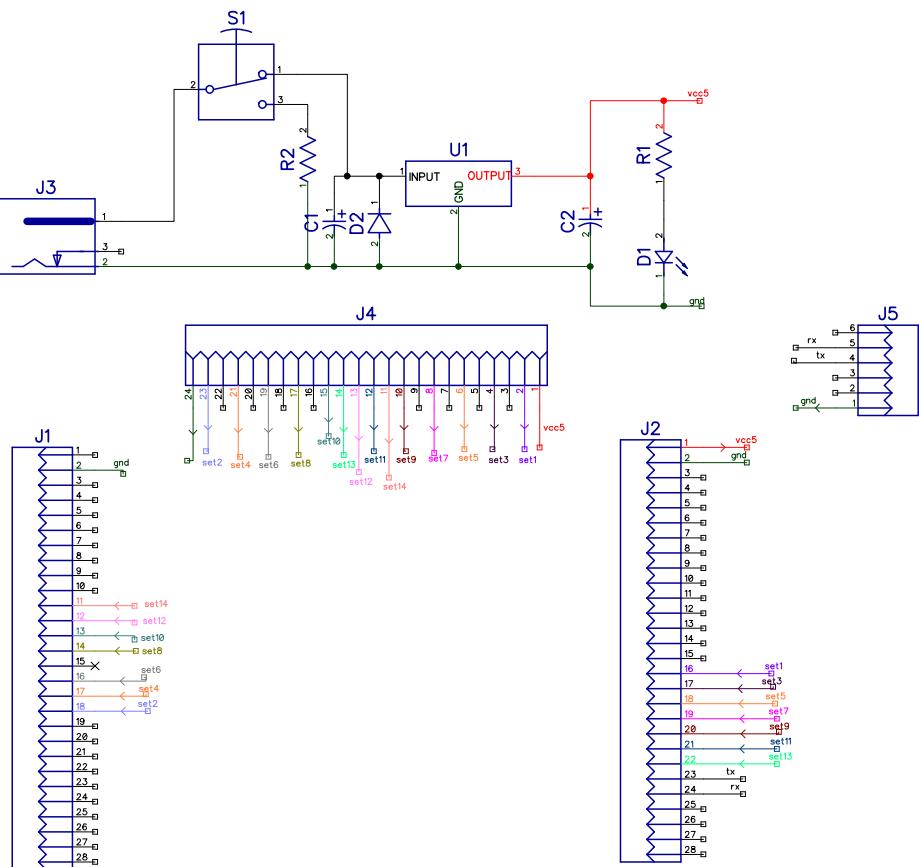


Fig. B.1 Schematic diagram of the discovery board support

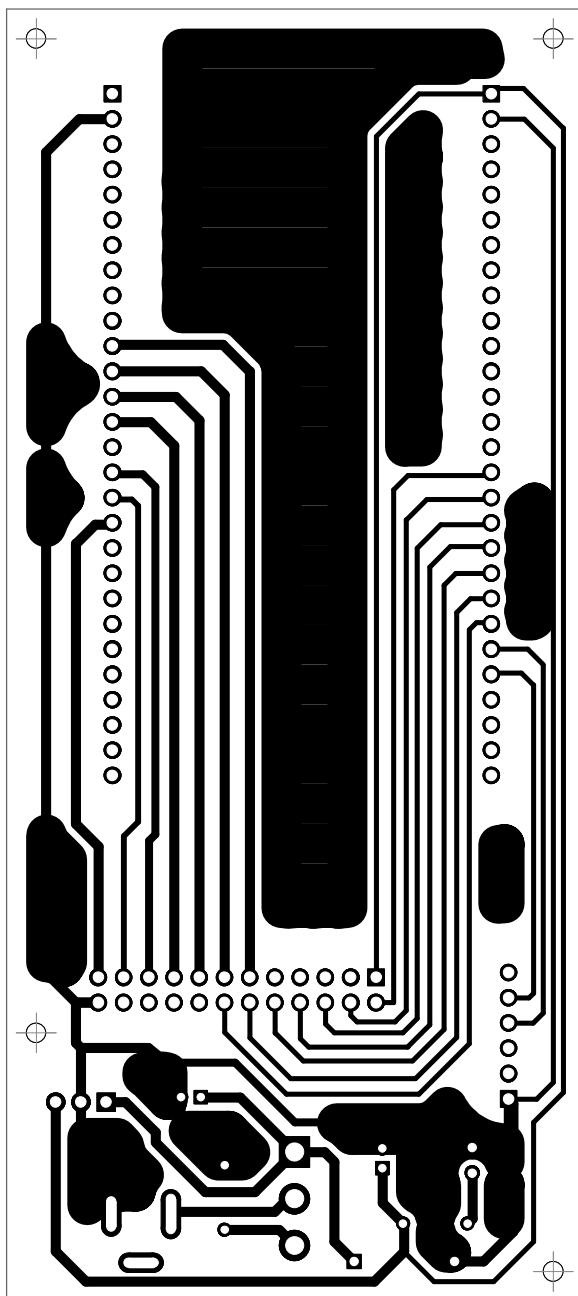


Fig. B.2 Layout of the discovery board support