AN AUTOMATED DNA STRANDS DETECTION SYSTEM FEATURING 32-BIT ARM7TDMI MICROCONTROLLER AND VGA-CMOS DIGITAL IMAGE SENSOR

by

MOHD HALIM MOHD NOOR

Thesis submitted in fulfillment of the requirements for the degree of Master of Science

UNIVERSITI SAINS MALAYSIA

March 2009

ACKNOWLEDGEMENT

My humble gratitude towards almighty Allah S.W.T for blessing me with opportunities and guidance. I would like to thank my wife, Zuraidah Sukadi for your prayers, patient and love. My two children, Amna Hanina and Mohamad Amirin Hafif, I am sorry to have sometimes neglected both of you to pursue my dream. To my dad and mom, thank you for the support and prayers – this thesis is for both of you. To my father and mother-in-law, thank you for always being understanding all this time.

I would like to dedicate my special thanks to my supervisor Professor Dr. Mohd Zaid Abdullah for his constant supervision and guidance. He always provided constructive criticism and detailed comments throughout my studies until I managed to complete this thesis. His efforts will always be remembered and appreciated.

I would like to thank Professor Dr. Mohd Nazalan Mohd Najimudin and Associate Professor Dr. Mohd. Razip Samian from School of Biological Science for their useful advice and aid during the experimental work at Biotechnology laboratory. I would also like to thank the staff of School of Electrical and Electronics Engineering of USM for their kind cooperation especially the laboratory technicians, En. Amir Hamid, En. Nor Azhar Zabidin and En. Aswadi Mohd. Desa.

I would like to thank my course mates and friends without whose support, help and friendship things would have difficult. My sincere thanks to USM for providing the necessary facilities and opportunities.

Mohd Halim Mohd Noor

March 2009

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AHB: Advanced High Performance Bus

APB: Advanced Peripheral Bus

CCD: Charge-coupled Devices

CE: Chip Enable

CFA: Color-Filter Array

CMOS: Complementary Metal Oxide Semiconductor

CPSR: Current Program Status Register

CPU: Central Processing Unit

DNA: Deoxyribonucleic Acid

DRAM: Dynamic Random Access Memory

EEPROM: Electrically Erasable Programmable Read-Only Memory

FPGA: Field Programmable Gate Array

GPIO: General Purpose Input/Output

GUI: Graphical User Interface

I^2C: Inter-Integrated Circuit

IPC: Industrial PC

LR: Link Register

OCT: Optical Coherence Tomography

OE: Output Enable

PC: Program Counter

PDF: Probability Density Function

PIC: Programmable Interface Controller

PIXCLK: Pixel Clock

PNG: Portable Network Graphics

RISC: Reduced Instruction Set Computer

SCL: Serial Clock Line

SDA: Serial Data Line

SP: Stack Pointer

SPSR: Saved Program Status Register

UART: Universal Asynchronous Receiver/Transmitter

VFM: Video Frame Memory

WE: Write Enable

LIST OF PUBLICATIONS & CONFERENCES

M.H. Mohd Noor, M.Z. Abdullah, “A Portable DNA Strands Detection System Featuring 32-bit ARM7TDMI Microcontroller and VGA-CMOS Digital Image Sensor” International Conference on Robotics, Vision, Information and Signal Processing, Nov’07, Malaysia.

AN AUTOMATED DNA STRANDS DETECTION SYSTEM FEATURING 32-BIT ARM7TDMI MIRCONTROLLER AND VGA-CMOS DIGITAL IMAGE SENSOR

ABSTRACT

Genetic DNA recognition is a routine experiment for detecting the origin of the species. Electrophoresis is one of the processes for such detection which has been used extensively. One popular technique based on electrophoresis is agarose gel sequencing. This method separates DNA molecules by size and eventually produces high resolution DNA strands. It is achieved by moving the negatively charged molecules through an agarose matrix towards cathode and anode if the molecules are positively charged. Manual analysis on the agarose gel image, usually is time consuming and prone to human errors. Automated agarose gel image analysis nullifies the problems and speed up the process. Hence the objective of this project is to develop an imaging system for automated DNA strands detection. This project can be divided into two stages. The first stage is the development of an imaging system that is able to capture agarose gel electrophoresis images. An electronic circuit is designed to interface and control two main components of the imaging system which are CMOS image sensor and microcontroller. The design is focused on developing a memory system that can temporarily store the image data before it is used for image processing. The imaging system was tested on a real electrophoresis application. The system is successfully able to capture agarose gel electrophoresis image of size 320×240 of 10-bit length. It was observed that the images were heavily corrupted with noise inherently present in the imaging system. The second stage is the development of image processing algorithms and pattern recognition software. The DNA detection algorithm adopts image subtraction technique to establish DNA recognition. For simplicity the algorithms are implemented in MATLAB® environment. The image comparison algorithm is based on image subtraction technique. The reference image is subtracted with the sample image and the maximum intensity value and average of intensity value of the resultant image are determined. The values are compared against the threshold values to establish recognition. Twenty samples of five types of DNA have been used to evaluate the accuracy and fidelity of the proposed procedures and methods. It was found that the system is able to classify DNA strands with 98.16% of average accuracy rate.

SISTEM MENGESAN LEMBAR DNA AUTOMATIK MENAMPILKAN MIKROPENGAWAL ARM7TDMI 32-BIT DAN PENDERIA IMEJ DIGITAL CMOS VGA

ABSTRAK

Pengecaman genetik DNA ialah eksperimen rutin untuk mengesan asal usul sesuatu spesis. Proses electrophoresis ialah salah satu proses pengecaman yang digunakan secara meluas. Jujukan gel agarose sequencing ialah teknik popular yang berasaskan Electrophoresis. Teknik ini mengasingkan molekul-molekul DNA mengikut saiz dan seterusnya menghasilkan lembar DNA yang beresolusi tinggi. Ia terhasil dengan menggerakkan molekul-molekul yang bercas negatif melalui gel agarose ke arah katod dan sebaliknya. Penganalisaan imej gel agarose secara manual mengambil masa yang lama dan cenderung kepada kesilapan. Masalah-masalah ini boleh diatasi dengan mengautomatikkan analisis gel agarose dan seterusnya mempercepatkan proses tersebut. Oleh itu objektif projek ini ialah membangunkan sebuah sistem pengimejan untuk mengesan lembar DNA secara automatik. Projek ini terbahagi kepada dua peringkat. Peringkat pertama ialah pembangunan sistem pengimejan yang mampu untuk mengambil imej gel agarose. Sebuah litar elektronik direkabentuk untuk mengantaramukakan penderia imej CMOS and mikropengawal. Fokus rekabentuk litar ialah mengarah kepada sistem memori yang mampu untuk menyimpan data imej secara sementara bagi pemprosesan imej. Sistem pengimejan diuji dengan aplikasi electrophoresis sebenar. Sistem tersebut berjaya mengambil imej gel agarose bersaiz 320×240. Imej yang diambil terjejas teruk dengan isyarat hingar yang wujud daripada sistem pengimejan. Peringkat kedua ialah pembangunan tatacara pemprosesan imej dan perisian pengecaman corak. Tatacara mengesan DNA menggunakan teknik pembandingan imej bagi menentukan pengecaman DNA. Bagi

memudahkan pelaksanaan tatacara pemprosesan imej dilaksanakan menggunakan perisian MATLAB®. Tatacara pemprosesan imej adalah berasaskan teknik pembandingan imej. Rujukan imej dibandingkan dengan sampel imej. Maksimum keterangan piksel dan purata keterangan piksel imej hasil pembandingan dikira. Jumlah yang dikira tersebut dibandingkan dengan jumlah ambang bagi menentukan pengecaman. Dua puluh sampel untuk lima jenis DNA digunakan di dalam menguji ketepatan prosedur dan kaedah yang dicadangkan. Hasil ujian menunjukkan sistem yang dibangunkan berjaya mengklasifikasikan lembaran DNA dengan purata kadar ketepatan 98.16%.

CHAPTER 1 INTRODUCTION

1.0 Overview of Computer Vision

Computer vision field has been researched actively since the last two decades. A research paper on this topic can be found back in 1975 which discussed the development of the related hardware for computer vision (Chien and Snyder, 1975). Among the many computer vision fields, image processing and image recognition are the two areas which are widely used in daily life application. It is proven by the widely used of applications such as bill identification, bar code reader in stores (Youssef and Salem, 2007) and fingerprint recognition for personal identification (Greenberg et al., 2002). The realization of all these applications demands complex algorithms and computation and needs to be processed by a powerful hardware system. But nowadays this is achievable due to the fact of technological advancement of software and hardware for computer vision in terms of processing power and speed.

Automation has been making a significant impact into the human daily life since its introduction. Nowadays automation can be seen in many applications such as industrial machinery, processes, manufacturing etc. It plays an important role and offers many advantages in human daily experience. Most notably, automation minimizes human errors and speed-up the rate of production. How does automation made such a significant impact in human daily life? Dictionary defines automation as operation of a process, apparatus or system by mechanical or electronic devices that take the place of human labor. Referring to the dictionary's definition, automation reduces the need for human intervention in the field of work. Machine and device handled by human operator assisted them in works that involves physical requirement.

DNA detection is a routine experiment for detecting the origin of the species. Two processes which are important for such detection are the electrophoresis and image formation. One popular electrophoresis techniques which is regularly used in genetic finger-printing is based on Agarose gel electrophoresis. Agarose gel electrophoresis is a technique that produces high resolution DNA strands which is then to be analysed to determine the DNA origin. Traditionally the analysis of DNA strands is repetitive and requires a lot of time (Kaabouch and Schultz, 2007). But nowadays there are numerous softwares available in the market to assist a researcher in the analysis of DNA strands image. These software performed analysis on the DNA strands image semi-automatically which is more efficient compared to the manual method.

1.1 Research Problem

Gel electrophoresis is a tool for gene and genomic analysis. The technique has been extensively used for gene identification, isolation and purification (Lin et al., 2007). Gel electrophoresis technique produces images consist of several lanes with a lane corresponds to a sample. Each lane has numerous bands and these bands contain valuable information to scientist. The information lies in the position of each horizontal band in a lane. The origin of a sample can be determined by comparing it to a known sample. Two samples are considered to have the same origin if they produce the same pattern of strand. The process of extracting this information manually consumes a lot of time and repetitious. In the worst scenario, the quality of the image produced is bad and inspection cannot be done. Thus in this case no information can be extracted and these lead to a waste of time and cost. In addition scientists have spent a lot of time to prepare the samples in the process and no results are acquired.

There are many commercial softwares have been developed for gel image analysis such as LabImage™, TotalLab™ and Gel-Pro Analyzer™. But these softwares have a few drawbacks such as semi-automatic techniques and require the manual setting during the gel image analysis. The adjustment parameters are not fixed for all bands. This means that the settings need to be continuously updated for a new bands (Kaabouch and Schultz, 2007). Here, the process becomes time consuming, make worst by the need for scientist to determine the best setting in addition to constant changing of parameters.

The most efficient approach for improving these circumstances is implementing all the tasks automatically. Automated analysis and detection can overcome the drawbacks that the commercial software are suffering and the manually information extraction and detection processes by scientist. The important processes of gel analysis are image enhancement, restoration and detection. The advancement of image processing algorithm would made automated analysis achievable. It is just a matter of combining the algorithms to obtain the best solution for automated gel analysis which can produce the most accurate result.

At present the DNA finger-printing can only be carried-out in the laboratory because the apparatus needed to form the analysis are not easily portable. This limits the use of this technique especially in field applications where the analysis has to be performed in-situ. A portable DNA inspector is thus needed for this specialised application. Hence the objective of this work is to design and develop a portable inspector system for rapid on-line analysis of DNA strands.

1.2 Objective

There are three main problems associated with the presence technique in agarose gel electrophoresis. First, the technique is done manually and thus consuming a lot of time. In few cases the quality of the image produced is bad and inspection cannot be done. As a result no information can be extracted. Second, the agarose gel electrophoresis can only be carried-out in the laboratory. This limits the use of this technique especially in field applications where the analysis has to be performed in-situ. Third, the analysis softwares for agarose gel electrophoresis are semi-automatic and hence requiring constant adjustment and setting. Therefore the objectives of this research are:

1. To design a customize image capturing system for DNA strand detection based on ARM7TDMI microcontroller and VGA CMOS image.
2. To develop a robust algorithms for image enhancement and DNA detection.

1.3 Research Methodology

The vision system mainly consists of a camera and a controller. The camera will provide image information and the controller will control the operation of the camera and the system. In order to avoid any complexity and difficulty in the implementation of image processing algorithm, the camera must be able to acquire quality images. The images require image processing algorithms to detect the DNA's origins. The development has been divided into two main stages. The first stage focuses the hardware development which includes electronic circuit interfacing and wiring up CMOS image sensor. The second stage is the software development which includes coding the algorithm for DNA detection and recognition.

1.4 Thesis Outline

The thesis chapters have been organized into six chapters as follows. The literature review has been provided in the Chapter 2. This chapter begins by presenting reviews of various imaging system that has been developed and used for application related to imaging purposes. This chapter also discusses several common microcontrollers that have been used in developing a system. The discussion outlines their usage and peripherals that have been used in the developed systems. Next the chapter presents the image sensor available in the market and their advantages and disadvantages. The following section details the image processing techniques that are suitable for the purpose of image enhancement and DNA detection. The theory chapter covers the basic theory of related hardwares and softwares in this project.

The next chapter details the hardware design aspects for this project. Details related to the microcontroller and image sensor used are presented in this chapter. The later section details the experimental setup and procedures that have been followed in this project.

The results chapter provides the results that have been obtained from following the steps of the project. The chapter also outlines the details of image processing technique that have been performed in this project. The accuracy of image processing algorithm has been discussed in the last section of this chapter.

The final chapter provides an overview of the objectives achieved by this project work. It also provides the proposed expansion that can be realized for this project.

CHAPTER 2 LITERATURE REVIEW

2.0 Introduction

This chapter provides an extensive review of literature relevant to the research addressed in this project. The important fields relevant to this research are imaging system for automated DNA detection featuring ARM7TDMI microcontroller and MT9V011 CMOS image sensor. The most important aspect of developing a DNA imaging system is the capability of the system to capture images and image processing procedure of agarose gel electrophoresis image. Various image processing techniques need to be applied to gel image before it can be used in the later stage which is DNA detection. DNA detection is an algorithm which is used to detect the position of bands and to perform pattern recognition and classification.

Gel electrophoresis is a basic analytical tool for biologist and has been extensively used in their researches and works. In biological applications, several dozen electrophoresis tests are performed each day which made an automated analysis of gel electrophoresis image is highly desired (Costa, 1999). Traditionally, the gel image needs to be analysed first before information can be extracted from it. But the analysis has many shortcoming particularly information extraction. Typically the analysis works usually involves many steps hence requires a lot of time. If the analysis procedures can be automated or even semi-automated, the repetitive works could be avoided hence it would save the researcher in terms of times and costs. In addition to speed-up the analysis, computerized analysis is reproducible and not prone to errors (Adiga et al., 2001).

2.1 Imaging System

A CCD-based system is designed for DNA detection in electrophoresis gels by UV absorption (Mahon et al., 1999). The technique used to detect DNA is based on ultraviolet (UV) absorption by nucleotides. A deuterium source illuminates individual sample lanes of an electrophoresis gel via an array of optical fibres. As DNA bands pass through the illuminated region of the gel the amount of UV light transmitted is reduced because of absorption by the DNA. During electrophoresis the regions of DNA are detected on-line using a UV-sensitive charge coupled device (CCD). The result shows that the system has lower possible limit of detection compared to previous works.

A CCD-based system is designed for determination of DNA in Agarose gel with array detection (Lim and Pardue, 1996). The experimental procedures involve gel electrophoresis with ethidium bromide staining followed by DNase treatment of the agarose gel containing the single DNA band. Then a CCD camera is used to monitor the fluorescence of the DNA band. The CCD camera (Thompson TH 7895) used is connected to an electronic unit (Photometrics CE200A) which provides camera, temperature control and analog-to-digital signal processing. The image data is sent to a personal computer for image processing. The computer is also used for controlling the camera function through a software package (Photometrics PMIS 2.1). The result of the research shows that the advantages of predictive kinetic method for the determination of DNA in solution to DNA determination in electrophoretic gels.

An imaging system is designed for automatic analysis for DNA electrophoresis image (Haydar and Jaggi, 1989). The system employs a CCD video camera (Panasonic WV CD-24) as its imaging device, a frame grabber (Matrox PIP-

1024) and computer for image data storage and processor. The size of image captured is 512×480. The image enhancement techniques used in this research are thresholding and segmentation. The result shows the system is robust and successfully employed for DNA electrophoresis analysis.

A monochrome charge-coupled-device (CCD) video camera is interfaced with a personal computer to quantifying the distribution of DNA molecules in agarose gels after Pulsed-Field Gel Electrophoresis (Dewey et al., 1994). The image data is digitized using a frame grabber and analysed using a video analysis software (JAVA v1.40, Jandel Scientific, Corte Madera, CA) installed in a computer. The CCD camera captures an image and is stored in the frame grabber. The JAVA image-processing system is used to collect data from images displayed on the computer's monitor. The computer performs an analysis on the images and the results are produced with the Cricket Graph software.

A system based on a CCD camera is designed for automated analysis of electrophoresis gels by image digitalization and processing (Costa, 1999). The image digitalization system is formed by a non-intensified color CCD camera interfaced to a microcomputer through a frame grabber that allows image acquisition to be performed in real time. A few different macro routines are built for image processing and data treatment. The routines are designed to be able to cope with the different requirements, specifically for each of the most frequently used types of electrophoresis gels. The CCD camera captures the images and is digitized using a frame grabber. Then the images are enhanced by employing histogram equalization, contrast enhancement techniques, noise reducing by low pass Fourier filtering; edge enhancement and location by the application of differential operators, posterization or binarization; blob analysis, matching and feature extraction. The research shows image digitalization and processing techniques possible to be used for electrophoresis gels analysis.

A computer based analyzer for gel electrophoresis images is designed (Sixiang et al., 2002). The instruments used in this system are UV light source, digital camera and camera obscura. The camera obscura is used to minimise interferance from external light source and to protect the system operator from the UV light. The digital camera is embedded with a memory system so as image data can be stored after image capturing. The digital camera is interfaced with a computer using USB port. The system can be operated to save, process, analyse image and compile and print an analysis report.

A paper presents an alternative scheme of detecting low intensity images with a CCD camera (Kolbe and Turko, 1989). The technique is illustrated by capturing images of ethidium bromide stained DNA electrophoresis gel using an uncooled CCD camera. The result shows that with suitable averaging, video images can be acquired using an uncooled CCD camera.

An embedded system is developed for an eye-detection sensor (Amir et al., 2005). The system is implemented using simple logic gates with no CPU and no addressable frame buffers. The devices used in the system are the Zoran 1.3 Mega pixel CMOS image sensor and the Altera Apex20k FPGA for the image processing. The FPGA was programmed to have a FIFO frame buffer to store the stream of video data from the CMOS image sensor, four components called frame subtraction and threshold, line, connect and centroid calculation. These components performed computation for pupil detection and the output was in the form of XY location within the frame region. The output of the system was transferred to a PC computer through USB interface for display and verification.

A real-time detecting system based on the DSP processor is developed for driver's fatigue (Jie et al., 2006). The system used real-time operating system DSP/BIOS to build tasks and interruptions, and will give out an incurrence of phonetic warning when the driver is tired. The system consisted of TMS320C6416 DSP processor and a CCD camera. Two parts of human face were used in driver's fatigue detection which were mouth and eye. Another system was developed for visually monitors the face of a driver and finds the location of the driver's eyes in real time (Betke and Mullally, 2000). The system addresses the actual environment within the interior of a moving car. The system used Sony DCR-TRV510 color video CCD camera which mounted on the dashboard. A 450 MHz dual processor PC with 384 MB RAM equipped with Matrox Meteor II image capture board was used to process the video.

A paper by Yeung et al., (2005) describes an embedded multi-user login system for fingerprint recognition. The system is divided into two parts: registration and verification were built using the Sitsang development board with Intel® PXA255 processor. The system operations were fingerprint acquisition, preprocessing, minutia extraction, match, identification, user registration, and template encryption. The fingerprint verification was fully tested on a fingerprint database consists of 1149 fingerprint images. The result shows the system achieved an accuracy of more than 95%.

A door-opening system is developed using a low-cost fingerprint scanner and a PC (Faundez-Zanuy, 2004). The system used a PC with Intel® Pentium III 800MHz as the main controller for the system. A special program was written in visual C++ to store and process fingerprints. Microsoft® Access was used to create the database of the fingerprints. The system also recorded the user names who accessed the door along with the access time for security monitoring purpose.

An authentication system based on fingerprint recognition is developed to improve the security protection of the mobile phone (Su et al., 2005). The authentication system was composed of two parts. One was the front-end fingerprint capture sub-system and the other was back-end fingerprint recognition sub-system. The fingerprint capture sub-system was an external module which consisted of LPC2106 ARM® processor and AT77C101B thermal fingerprint sensor. The fingerprint recognition sub-system was based on the mobile phone. The recognition sub-system was a program that performed fingerprint registration and matching. The fingerprint capture sub-system was interfaced with the fingerprint recognition sub-system via UART.

An imaging system is developed for traffic measurement (Nishiyama et al., 1991). The imaging system measured the traffic flow and provides information to the central computer regarding the traffic volume, mean velocity and congestion rate. The system consisted of 3 sections: the cameras, the pre-processor and the main processor. The pre-processor was composed of slave CPU (S-CPU) which was the Motorola MC68000 with 12.5 MHz and Video Frame Memory (VFM). The Motorola MC68000 with 10 MHz was used for the main CPUs (M-CPU). The camera captured video and sent it to the VFM. This device performed A/D conversion process and stored the image signals in the form of 8-bit length data. The data was processed by the S-CPU at a high speed to detect vehicles which assumed to be stationary. The information was sent to the M-CPU. Here, the data were processed using the measurement algorithm, from which the illegally parked vehicles were identified.

A machine vision system is developed for sorting sweet tamarind (Jarimopas and Jaisin, 2008). The system used charge-coupled devices (CCD) camera connected to frame grabber card slotted in a microcomputer. Two microcontrollers, ATMEL AT89S8252 as the master and AT89C2051 as the slave were used in the project for electronic control unit. When a photosensor comprised of laser diode and photo transistor detected a sweet tamarind on the conveyor, a signal produced by the photosensor was sent to the master microcontroller. The signal will be processed and sent to the frame grabber in the microcomputer through a serial port. The microcontroller commanded the CCD camera to capture an image and the image data was sent back to the frame grabber. The microcomputer processed the digital image for size, shape and defect to produce a result and the information was sent to the master microcontroller to be converted into codes of grades. The information will be compared by the slave microcontroller with the predetermined grade for sweet tamarind identification.

A fuzzy logic control of a multi-spectral imaging sensor is developed for in-field plant (Kim et al., 2008). A multi-spectral imaging sensor was used to estimate the plant health responses such as plant nutrient, water content, leaf senescence and plant response to nitrogen by assessing the spectral of plant leaves reflectance by solar radiation. The spectral assessment needs to be made under consistent illumination but the illumination is influenced by the solar radiation changes. The authors developed a control algorithm that could obtain a consistent image quality using fuzzy logic under the variation of illumination. The system consisted of a multi-spectral imaging sensor (MSIS), a portable computer equipped with frame grabber, an ambient illumination (AI) sensor and a GPS. The MSIS captured images with a resolution of 640×480 at 8-bit/pixels and the ambient illumination is recorded by the AI sensor. The gain and exposure of the MSIS were adjusted by referring to the recorded ambient illumination by the AI sensor to centralize the grey-level distribution of images. Spray nozzle was activated based on the result of estimation of plant leaves reflectance.

2.2 Processor for Imaging System

Image processing of DNA detection is not a real-time critical application where the output is dependent on the duration of the algorithm computation. In this project, the algorithm computation will not be carried out by the processor. There are two main tasks of the processor in this project. One is to control the image sensor and the overall operation of the system. The other is to send the image data to the computer. Therefore a simple and general purpose microcontroller is sufficient enough to be used in this project development and powerful processor such as DSP and FPGA are not considered. These powerful processors are not only expensive but the implementation would be unnecessarily complicated.

Although the processor criterias are not demanding, there are a few attributes that are needed to be considered. The microcontroller must have UART and I^2C peripherals for intefacing with the computer and CMOS image sensor. Preferably must have sufficient general purpose pins for controlling the system's components such as memory, counters and etc. The other consideration is the microcontroller must be 16-bit or more since the CMOS image sensor is a 10-bit device.

The 8051 microcontroller was introduced by Intel in 1980 as an improvement over its predecessor 8048 microcontroller. It was widely used and one of the most popular in 80s and 90s. The typical features of 8051 microcontroller are 4 Kb program memory, programmable I/O ports, timers, counters and serial data communication. The 8051 is also a low-power consumption microcontroller hence it is suitable for battery-powered device. Over the years, companies such as Infineon Technologies, Maxim Integrated Products via its subsidiary company Dallas Semiconductor, NXP formerly known as Philips Semiconductor, Winbond, ST Microelectronics, Silicon Laboratories formerly known Cygnal, Texas Instruments and Cypress Semiconductor have manufactured and provided a broad range of 8051 based microcontroller. The 8051 microcontroller has been used in numerous applications such as keyboard, display device such as seven-segment numeric display and LCD display, pulse measurement and A/D and D/A converter (Ayala, 1991).

Programable Interface Controller (PIC®) is a Harvard architecture based microcontroller developed by Microchip Technology Inc. Typically, PIC® is a 8-bit microcontroller having RAM data memory, ROM or flash program memory, digital I/O ports, timers and external clock interface. The 16-bit architecture is available which are faster and with more memory capacity for data and program and 32-bit architecture is in production and will be available in the future. The 16-bit PIC® microcontroller can perform up to 40 MIPS and is featuring more complex peripherals such as I^2C, Universal Serial Bus, Serial Peripheral Interface, Controller Area Network and Infrared Data Association. PIC® is one of the most popular microcontroller among the developers and hobbyist due to its low cost and wide range of availability. Until February 2008, Microchip has managed to sell up to 6 billion PIC®s (Lawson, 2008). PIC® can be used in applications such as automotive, consumer electronics, medical applications and telecommunication.

PIC16F87 had been used to develop an interface circuit for micromachined gas sensors (Baroncini et al., 2003). The PIC® was interfaced with a gas sensor through a data acquisition circuit by means of built-in ADC. UART module of the PIC® was interfaced with a computer to allow external input from the user to control the operation of the system and receive the sensor output value. A paper written by Bayindir and Ates, (2007) demonstrated the development of control unit for a friction welding machine using PIC16F877 microcontroller. The PIC® was interfaced with 2x16 LCD modules, keypad and a circuit of motors and valves. The PIC® was preferred because the main components such as RAM, EEPROM and peripherals are contained in the same integrated circuit. Thus the size of the developed system would be decreased, the cost can be reduced and the designed of the system became easier. In addition to those mentioned, PIC® is a CMOS based chip thus it is suitable for development of low-power consumption devices.

ARM® microcontroller was introduced by ARM Limited in 1985. The ARM® is the most widely used 16/32-bit RISC processor. In the last decade, this processor captures approximately 75 percent of the RISC processor market due to its advantages such as low cost and energy consumption and high performance (Zhang et al., 2007). The ARM® features such as Thumb instruction set which requires 16-bit wide data bus, increases the ARM® capability and flexibility. The Jazelle technology allows java application to run on ARM® processor.

Sung et al., (2002) had demonstrated the used of Samsung S3C44B0X ARM7TDMI processor in the development of multimedia handheld device. The ARM7TDMI was chosen because of its good compiler support, low cost and its system integration. The project highlighted the ARM7TDMI capability of implementing the DSP algorithm despite of lacking programmable DSP features. The implementation was achieved by exploiting the ARM7TDMI processor designed features such as ARM7TDMI large number of internal registers and 32-bit barrel shifter.

Samsung S3C44B0 ARM7TDMI processor had also been used in the development of autolevelling control system of carding machine (Xu et al., 2006). The ARM7TDMI was used to replace the usually used Industrial PC (IPC) in carding machine. In order to achieve this objective, the ARM7TDMI functionality was expanded by interfacing the ARM7TDMI with two boards which are memory system and peripheral devices. The memory system board consisted of DRAM for main memory application, flash memory for application program and EEPROM for control parameters and user information data files. While the peripheral devices board consisted of serial driver chip, high speed counter, A/D converter and I/O port control.

2.3 Image Sensor

One of the main components of a digital imaging system is an image sensor. It is widely used in digital cameras and other imaging devices. An image sensor is a device that converts the light from a scene that incident at its surface into an array of electrical signals. Color imaging is achieved by using a color-filter array (CFA), typically green-red-green-blue Bayer CFA, which is placed on top of the image sensor pixel array. By using CFA, signals that are produced will emit one of the three colors (Gamal and Eltoukhy, 2005).

There are two possible technologies of image sensor which are complementary metal oxide semiconductor (CMOS) image sensor and charge-coupled devices (CCD) image sensor. CMOS image sensor has been introduced around 1960s and it used photodiode image sensor with MOS scanning circuit. It was the first solid-state imager that had been invented. However the technology was not well received due to its poor performance and large pixel size. In 1970, CCD was invented and quickly became the dominant image sensor because it offers superior performance in terms of dynamic range, low noise, smaller pixel and high sensitivity to light (Bigas et al., 2006).

In the early 1990s, the passive pixel sensor (PPS) technology was introduced and adopted the CMOS image sensor. This technology is an improvement from the earlier technology, however it still offered much lower performance compared to CCD. Further research was done and active pixel sensor (APS) was subsequently introduced. The APS is a major improvement to the PPS in many aspects. The former offers higher speed readout and high signal-to-noise ratio (SNR). Hence CMOS image sensor has emerged as an alternative to CCDs (Bigas et al., 2006).

Despite the facts that at the moment CMOS image sensor can compete with the CCD technology, it is still behind in a few significant aspects. These disadvantages are less sensitivity to incident light, prone to different noise sources, less dynamic range and less image quality than CCD. Knowing these shortcomings, research on CMOS image sensor mostly focused on overcoming those disadvantages. However there are several advantages of CMOS image sensor over CCD's technology which can be highlighted. The advantages are:

1. Low power consumption
2. Lower cost compared to CCD's technology
3. Compatibility with standard CMOS technology
4. Small size
5. Random access of image data
6. No blooming and smearing effects
7. Simple to design with

Among the advantages which have been outlined, the most significant advantages of CMOS image sensor is the relatively low cost and the ability to integrate and interface with standard CMOS technology.

2.4 Image Enhancement

The most important aspect of automated analysis of gel electrophoresis image is an efficient and accurate image processing and detection algorithms. DNA detection can be successfully performed if the gel image is good in quality despite the fact that the main content of the image are the DNA bands and the wells. The context of good quality is the band can be distinguished from the background. Usually the captured gel image is corrupted with noises and has non-uniform background. Although the bands are visible to the naked eyes, but frequently some of the bands are faint or washed out due to the fact that their gray levels are about the same as the level of the background. In addition to that, sometimes the bands have smear effects which are produced during the electrophoresis process.

The complexity of DNA detection is influenced by the result of image processing procedure. DNA detection can be very straight forward if the quality of the image is high in which free of noise and it has a uniform background. For example if the image is in binary form, the pattern of the DNA strand can easily be recognized just by detecting the white pixels in the image. But to convert the image into a binary form is not an easy task due to the factors that have been mentioned. Basically the image analysis of gel image is about separating the background from the foreground. There are several methods of processing the gel image in order to obtain an image that can be used for detection. The general methods are contrast enhancement and noise filtering.

The approach of contrast enhancement is a transformation of an image to another image with higher contrast between the white and black pixels. Some of the basic contrast enhancement techniques are image negatives, log transformation and power-law transformation (Gonzalez and Woods, 2002). These techniques are the simplest form of transformation in which a pixel value is map to another pixel value. Image negatives is basically an inverse transformation function. It is best suited for enhancing an image which dark area is dominant. Log transformation is used to widen the range of narrow low gray-level while narrowing the higher gray-level. The purpose of power-law transformation is more or less the same as log transformation. But it is more versatile in the sense that the power value can be varied and various output images will be produced.

Various methods and techniques have been developed for contrast enhancement which is based on histogram manipulation. One of the most popular techniques is histogram equalization. It has become accepted as the standard technique for image contrast enhancement. The idea behind histogram equalization is to distribute pixels uniformly across the gray levels so as an image will have a large variety of gray tones. As a result the image will exhibit more details and has dynamic range. But histogram equalization becomes less effective when processing an image that are vary in contrast characteristics (Stark, 2000).

While histogram equalization automatically determines a transformation function that will produce an image with uniform histogram, histogram matching technique specifies the shape of histogram as its transformation function. The benefit of this technique is that the desired output is predictable and yet it is easy to implement (Gonzalez and Woods, 2002). However histogram matching still inherits the histogram equalization limitation.

Over the years histogram equalization has been used to develop better image contrast enhancement techniques. Adaptive histogram equalization was developed to cater the limitation of histogram equalization. The concept of adaptive histogram was the same but local histogram is used instead. Adaptive histogram equalization has been proved to be effective such as in medical imaging. Despite its effectiveness, adaptive histogram equalization is not easy to be implemented due to its computational complexity (Stark and Fritzgerald, 1996). There are various algorithm of adaptive histogram equalization such as contrast-limited adaptive histogram equalization, generalizations of histogram equalization and variable regions. Readers may refer to publication by Pizer et al., (1990), Stark, (2000) and Vossepoel et al., (1988) for detail explanation.

Yiheng et al., (2007) presented the used of least squares fitting and histogram matching for contrast enhancement of optical coherence tomography (OCT) images. This technique was based on histogram matching with least squares fitting. The normalized input image histogram was fitted using least squares fitting algorithm to a probability density function (PDF) such as Gaussian function. Other PDF can be used which was best suited the input image histogram. Then the output image histogram was simulated using the fitted input histogram. The output image was transformed from the output histogram. The result showed that the output image had much higher contrasts and exhibited more details clearly. The key factor for this technique to be effective was choosing the best PDF to fit the input histogram.

The approach called noise filtering can be categorised into two according to the type of noises, additive noise and periodic noise. Some of the basic techniques of additive noise which are based on mean, arithmetic mean, geometric mean, harmonic mean and contraharmonic mean filters. All these filters have successfully eliminated Gaussian and uniform noise. The contraharmonic filter is more suited for impulse noise. But these filters also reduce the details of image in the process due to the effect of blurring (Gonzalez and Woods, 2002).

Order-statistics filter can remove salt-and-pepper noise better than mean filter (Gonzalez and Woods, 2002). This filter is based on ordering result of the pixels in the filter area. Some of the order-statistic filters are median, max and min, midpoint and alpha-trimmed mean filters. The alpha-trimmed mean and median filters give better noise reduction than the max and min filters with alpha-trimmed mean filter performs better than the median filter in terms of smoothing effect. The midpoint filter works best for Gaussian and uniform noises.

The better method than the previously discussed filter is adaptive filter whose behavior evolved with the statistical characteristics of the image inside filter area (Gonzalez and Woods, 2002). The parameters that can be used in determining the filter behavior are mean and variance. For example, the performance of adaptive median filter is superior to the conventional median filter in the sense of retaining the image details in the process of removing noise. But the drawback is the complexity of the computation and implementation of this filter. Another example of adaptive filter is adaptive Wiener filter. This filter works by performing large amount of filtering in the region which has small variance to achieve maximum noise reduction. In the region which has large variance, the filter performs small amount of filtering to preserve the image edges (Pham, 2001).

An efficient algorithm was developed for impulse noise reduction from corrupted images (Luo, 2007). The algorithm consisted of two steps which are impulse noise detection and impulse noise cancellation. Two statistical characteristics were developed before impulse noise detection is employed which were increasingly ordered of pixel values and rank-ordered absolute differences (ROAD). The two statistics were combined and used for impulse noise detection operation. The impulse noise detection was based on fuzzy detection scheme. This technique determines how many pixel was corrupted in a range of 0 to 1. 0 indicates a pixel is noise-free while 1 indicates a pixel is heavily corrupted. This parameter was used to determine how much noise cancellation will be performed on a pixel.

2.5 Image Segmentation and DNA Detection

Traditionally image segmentation algorithm can be categorized into two. The first is an algorithm that is based on discontinuity of intensity value and the second category is based on similarity of intensity value. One of the fundamental approach of similarity segmentation is thresholding. Thresholding is one of the most popular approach in image segmentation. It is mainly used in application that demands quick and fast processing time in producing the result. In addition to the processing speed, thresholding is simple and intuitive in nature in its implementation.

Thresholding has been used extensively to extract information or objects from the background in various image processing applications. The implementation of thresholding depends on the histogram of an image. An image with bimodal histogram can be easily partition by placing single global threshold between the dark region and light region. This type of thresholding is usually used in highly controlled illumination such as industrial inspection application. A novel operation using automatic thresholding and shifting was developed to equalize the gray level of agarose gel image background (Kaabouch et al., 2007). Although the image had non-uniform background, thresholding was successfully applied. As opposed to the traditional method, histogram of the image was not used in determining the threshold value. Instead the threshold value was a function of maximum and minimum intensity value of the gel image with a weighting value.

Local thresholding (adaptive) was used in medical imaging application such as X-ray where the illumination is not consistent. The objective of the application was to outline the boundaries of heart ventricles in cardioangiograms. The image was subdivided into 49 regions with 50% overlap over each other image. Only the regions which had bimodal histogram indicating the presence of a boundary are processed and the result showed that thresholding successfully achieved the application objective (Gonzalez and Woods, 2002).

Contour Area Filtering algorithm which was a form of local thresholding was developed to be used in 2D electrophoresis gel image analysis (Kazhiyur-Mannar et al., 2006). This algorithm was applied with a priori knowledge of the size of the interested objects in the image. The algorithm worked by first selecting a pixel with intensity, I. Suppose the pixel was foreground, then all the neigboring pixels with intensity equals or darker than I should also be foreground. This checking was repeated for all the neighboring pixels and if the checking area was greater than the size of object, the area was considered as a background.

A paper described bands and lanes segmentation using matched filter technique and band skeleton based calculation (Lin et al., 2007). The matched filter was designed based on the shape and intensity distribution of the bands. In this case, the band profile was bell-shaped, thus Gaussian distribution was used for approximation. Three parameters were defined for lane segmentation TH_{LO} , TH_{HI} and TH_{Lane}, which were smallest number of bands in a lane, largest number of bands in a lane and smallest lane width respectively. Calculation started with counting the number of nonzero pixels on each column in the image to obtain the vertical projection curve. The threshold TH_{LO} was used as the horizontal line, to cut the curve in to several components. A component with width greater than TH_{Lane} was considered a lane. This process was iterated until the threshold reached TH_{HI}.

An algorithm was developed for automatic segmentation of DNA bands by Akbari and Albregtsen, (2004). The algorithm developed was based on three parameters which are variance (var), Mean-Variance-Ratio (MVR) and equivalent width. An average of intensity profile along y-axis was used in the process of segmentation. The algorithm worked by extracting informations from the intensity profile. The center of local minimums in the intensity profile which represented the bands in the image were estimated to compute the center of the bands. The var and MVR were used in classifying the correct bands from the intensity profile and the width of bands was computed by applying the equivalent width method. This method was accurate if the intensity profile had clean curve. Thus the image had to be pre-processed first before the algorithm could be applied. The result showed that the algorithm successfully segmented the DNA bands. But the algorithm sometimes failed in classifying all the bands if the var and MVR value were fixed. This was due to the different in quality of images. In some cases noise could be classified as false DNA bands.

2.6 Conclusion

Most of the imaging systems that have been reviewed are CCD based. The CCD camera is connected to a computer either using USB port or a frame grabber. The computer is used for controlling the CCD camera and performs the image enhancement and analysis of DNA electrophoresis gel. A system installed with a frame grabber is used for acquiring the image as well as storing the image data. Most