2018

Automated Optical Mark Recognition Scoring System for Multiple-choice Questions

Murtadha Alomran
Edith Cowan University

Automated Optical Mark Recognition Scoring System for Multiple-choice Questions

Thesis submitted in fulfilment of the requirements

for the degree of

Master of Engineering Science

Murtadha ALOMRAN

Douglas Chai, Principal Supervisor

Alexander Rassau, Associate Supervisor

School of Engineering

Edith Cowan University

2018

ABSTRACT

Multiple-choice questions are one of the questions commonly used in assessments. It is widely used because this type of examination can be an effective and reliable way to examine the level of student’s knowledge. So far, this type of examination can either be marked by hand or with specialised answer sheets and scanning equipment. There are specialised answer sheets and scanning equipment to mark multiple-choice questions automatically. However, these are expensive, specialised and restrictive answer sheets and optical mark recognition scanners.

This research aims to design and implement a multiple-choice answer sheet and a reliable image processing-based scoring system that can score printed answer sheets and send back the scored answer sheets to students automatically. The proposed method will allow users to print the proposed answer sheet and use a normal scanner and computer to perform the scoring. Moreover, while compiling the assessment results, it will annotate the answer sheets with feedback and send back to students via email.

Firstly, the proposed system requires class list and scanned answer sheets to perform scoring. Then, the software algorithm will first locate the finder pattern points and correct the answer sheet tilt. After recognising the finder pattern points, the algorithm processes student information in an aim to match the decoded student information with the class test. After accomplishing student information recognition, based on the key answer, the algorithm scores the answer area which contains 72 answer areas and five choices in each. In the last stage of the processing, the algorithm annotates the answer sheet and store the score in a spreadsheet.

There are three major research contributions. First, a new method to change multiple answers on the answer sheet. On the proposed answer sheet, students do not need to use pencil and eraser to change answers. Second, a new method to recognise student information on the answer sheet. The proposed system replaces the conventional method of encoding student information on answer sheet by introducing new method to decode student ID and utilising the count of student name characters. Lastly, a new fast method to provide a test result feedback. The proposed system

provides annotations on the answer sheet and sending the answer sheet to student email.

After experimenting the system, the system results indicated that the system meets the research objectives. Speed wise, the system scoring speed is 1.25 seconds without annotation and 2.25 seconds with annotation. The software algorithm proved to be robust to detect the finder pattern points when noise exists in the answer sheet margin and pen scribbles around the finder pattern area. In addition, the algorithm is able to correct scanning tilt up to 5-degree rotation. Furthermore, the algorithm is capable of recognising different shading styles as long the shading area covers at least 40% of the answer box. A case study was conducted in a real test situation to retrieve more results about the system. The outcomes of the case study are that the success rate of finder pattern recognition was 100%, the success rate of marked answer recognition was 97.7%, and the success rate of student information recognition was 90.2%.

AUTHOR DECLARATION

I certify that this thesis does not, to the best of my knowledge and belief:

(i) *incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution of higher education;*
(ii) *contain any material previously published or written by another person except where due reference is made in the text; or*
(iii) *contain any defamatory material.*

Name: __________Murtadha Hussain ALOMRAN____________________________

Signature: [redacted] _ Date: ______02/11/2018_______

ACKNOWLEDGEMENT

I would like to express my deep gratitude to Edith Cowan University for providing their best facilities to all students and honouring me with a research scholarship. I am highly thankful to my project supervisor **Dr Douglas Chai** for his profound support and patience as he is highly busy working on his researches. Above all, I thank all my friends and my family specially my mother for all her prayers to me to achieve the outcome of this research.

The research case study was conducted with accordance with the Research Ethics Committee at Edith Cowan University. The research human ethics approval was received December 2017 from ECU Research Assessments and Student Records, Approval Number 19339.

ABSTRACT OF PUBLISHED PAPER

Citation

M. Alomran and D. Chai, “Automated Scoring System for Multiple Choice Test with Quick Feedback,” *International Journal of Information and Education Technology,* vol. 8, no. 8, pp. 538-545, Aug 2018.

Abstract

Although automatic scoring systems for multiple-choice questions already exist, they are still restrictive and use specialised and expensive tools. In this paper, an automated scoring system is proposed to reduce the cost and processing restrictions by taking advantage of image processing technology. The proposed method enables the user to print the answer sheets and subsequently scan them by an off-the-shelf scanner. In addition, a personal computer can process all the scanned sheets automatically. After scoring, the proposed system annotates the sheets with feedback and send them back to students via email. Moreover, two novel features are introduced. The first feature is the handwriting recognition method to recognise student ID. We called this the segmented handwritten character recognition. This new method replaces the conventional student ID recognition commonly known as the Matrix Identifier. The second feature is our specially designed answer sheet that allows students to change their answers with multiple attempts easily. As a result, there is no need to erase pencil shading or change the entire answer sheet if any mistake happened during the test. The proposed system is designed to be cheap and fast.

Table of Contents

1. INTRODUCTION

In the first chapter, an introduction of the research will be presented. Section 1.1 will introduce a brief background about the research area. The problem domain will be discussed in Section 1.2. Section 1.3 will state the objective of the research. Section 1.4 will explain the significance and research contributions. Finally, the thesis organisation will be mentioned in Section 1.5.

1.1 Introduction

In the need to improve learning efficiency, technology is an essential tool to improve the outcomes of educational systems. In the last two decades, technology has occupied a significant part of education sectors. Spending wise, most of education institutes budget is spent on education technologies such as online learning and computer-based education. For example, USA expenditure on education technology is expected to increase to be US$56 billion per year, which aims to improve education quality [1]. This indicates that technology is a key emphasis in education.

One of the education technology aspects that revolutionised education practices is assessment evaluation. The aim of the assessment is to measure student level of understanding. Assessment types can be generally categorised into two types: formative and summative assessments [2].

Formative assessments aim to measure student learning to give the assessors feedback such as presentations, and peer assessments while summative assessments aim to rate student learning on regular basis based on curriculum standards as an example final exams, mid-semester test, and quizzes [2]. Summative assessment is normally designed to assess students through a test paper. This kind of assessment requires long time for scoring especially with a large number of students.

As most assessments are paper-based tests, the question types of a test paper are many such as short answer, fill in the blank, true/false, matching, and multiple-choice questions. Multiple-choice question is one of the most effective technique to reduce scoring time. The multiple question concept is that students have to choose one answer out of multiple-choices provided. A multiple-choice question contains only one correct answer (usually) and wrong answers as distractors [3]. Multiple-choice test has multiple advantages that makes it preferable for assessors. The advantages of multiple-choice test can be listed as follows:

- quickly test student learning
- be applicable for a large number of students
- be quickly and automatically marked [4].

According to Pearson Education study [5], school teachers in Australia spend about 8 hours per week working outside the school preparing for school. One part of the preparation is scoring student test papers. The study assured that there is a need for greater learning technology adoption in schools to reduce the working hours outside the school time. Nowadays, there are some learning technologies have been adopted to make assessment scoring even easier and quicker such as offline computer-based test, online test, and a multiple-choice Test recognition system. However, there is still debates about the use of computers over the traditional multiple-choice test paper for assessments due to the cost of the computers to run tests for a large number of students (for example greater than 100 students) and reliability of computer during the test to hold a test without hardware or software issues [6]. These points make the traditional multiple-choice test more reliable although computer-based test still exists in most educational institutes but for a small number of students. Therefore, this raises the attention to improve the traditional multiple-choice test by taking advantage of computer technology and in the same way keeping the paper-based style. This can be achieved by implementing a reliable multiple-choice test recognition system.

A multiple-choice test recognition system offers the ability to mark a large number of test papers quickly. The idea of an automated multiple-choice test is to recognise a customised answer sheet and then compare it with the key answer to obtain the total mark. An optical mark recognition system is by far the trade-off to score a large number of exam papers due to the computer-based disadvantages [7].

1.2 Problem Domain

The existing optical mark recognition systems are expensive and restrictive. They require specialised scanners and answer sheets, which are expensive to afford specially for small institutes. Apart from the specialised scanners and answer sheets, the optical mark recognition systems put further restrictions such as:

- the colour of the pen
- the way to answer a multiple-choice question (shading)
- the inability for students to change their answers on answer sheet
- The way to encode student name and identification numbers.

All these restrictions add more complexity to the usability and decrease the flexibility.

1.3 Objectives

The objectives of this research are to:

- conduct an extensive literature review on assessment and automated multiple-choice scoring systems and explore the state of the art;
- design and implement an answer sheet that is user friendly, and improve the existing answer sheet features with many multiple-choice questions in one sheet;
- design and implement a fast and robust automated image processing based scoring system that can recognise and process the designed answer sheet and provide a quick feedback to students via email;
- test and evaluate the effectiveness, robustness, and speed of the scoring system.

1.4 Significance

The possibility of scoring multiple answer sheets of a large number of students can save a lot of time and reduce the cost of money and human resource for test scoring. Furthermore, the proposed optical mark recognition algorithm focuses on decreasing the restrictions on the students and the assessors. All these significances contribute to improve the quality of assessment evaluation. The contributions of the research are as follow:

- A new method to change multiple answers on the answer sheet: On the proposed answer sheet, students do not need to use pencil and eraser to change answers.
- A new method to recognise student information on the answer sheet: The proposed scoring system replaces the conventional method of encoding student information on answer sheet.
- A new fast method to provide a test result feedback: The proposed system provides annotations on the answer sheet and sending the answer sheet to student email.

1.5 Thesis Organisation

The thesis consists of five chapters as follow:

In Chapter 1, the thesis starts with a research introduction which includes the objective, thesis domain, and significance and research contribution.

Chapter 2 presents an introduction to assessment types and focuses on multiple-choice questions. Then, it reviews the automated multiple-choice test scoring solutions with the current state of the art. Lastly, the chapter explores image processing fundamentals as the proposed system is an image processing based.

Chapter 3 addresses deeply the methodology to score the answer sheet automatically. It starts with the overall system design. Then, the details of the answer sheet design and the scoring system are discussed. Lastly, the hardware and the software materials that are need to utilise the proposed methodology are presented.

Chapter 4 presents the result of the proposed methodology. Firstly, it presents the outcomes of answer sheet design. Then, the scoring algorithm is analysed to show the processing stages. Furthermore, the scoring algorithm robustness is discussed with results. Moreover, the speed of the scoring system is presented with a comparison with some of scoring methods. In addition, the chapter contains a case study of a real test situation. The chapter concludes with the proposed system limitations and a summary.

Chapter 5 presents the conclusion of the research. It starts by reviewing the research objectives and the outcome of the research. Then, it recommends the possible future work of the research area.

2. BACKGROUND

In this chapter, a review on multiple-choice assessment and scoring systems will be explored. The chapter starts with a review on assessment in Section 2.1. Section 2.2 carries on with a detailed background study on automated multiple-choice scoring systems. As the thesis is an image processing based, Image processing fundamentals and the key techniques will be covered in Section 2.3.

2.1 Assessment

Clegg and Cashin [8] argue that one of the institution responsibilities is to design classroom assessments, as it is one of the most significant parts of teaching and learning roles. According to Harlen and James [9], the goal of education is to attain learning and understanding together. The role of the teachers or lecturers is to teach students and make sure that students reach the appropriate level of understanding. This brought the need to assess students. Assessment can be defined as the tool to measure candidate level of understanding based on the curriculum criteria [10]. Educators benefit from assessments as they can evaluate student's understanding and progress. In general, the purposes of assessments are:

1. to promote learning experience
2. to provide a feedback about students' learning level
3. to help to improve students' learning
4. to provide a reliable record of students' achievements [11].

2.1.1 Types of Assessments

There are two general types of assessment, formative and summative assessment.

Ramesh and Sidhu [3] state that formative assessment can be defined as the assessment to help student to learn information and promote learning skills. Formative assessment may be conducted by the following activities; assignments within instruction materials, test or quizzes, and peers assessments.

Ramesh and Sidhu mention that summative assessment can be defined as the assessment to measure student outcomes of the learning materials. Summative assessment can be carried out via one of the following activities; Final examinations, course work, practical demonstrations.

The differences between the formative and summative assessment according to [12] are summarised as follow: Formative assessment is designed to test student learning at the beginning of educational period or during the progress of the semester whereas the summative assessment is designed to assess students at the end of the semester or instruction period. The formative assessment’s aim is to expose student reflect on the curriculum and monitor student progress during the period of the instruction. On the other hand, the aim of summative assessment is to expose student learning achievements via assessments that require from students to demonstrate their understanding of a particular unit. Usually, summative assessment is conducted during a class test or a final exam.

The types of questions vary. The most common question types are:

- long answer (essay)
- short answer
- matching
- true/false
- multiple-choice questions.

The essay question type is that students are required to write a long answer about the question topic. The answer usually is explanative and may contain several paragraphs. The short answer question style is that students are required to write between a sentence to a paragraphs explaining the question topic.

However, these two types of questions are tedious to score for assessors. Matching, true/false, and multiple-choice questions are quicker to score. Moreover, the score of these questions is either right or wrong which make them favourable for assessors in the test.

Matching questions require student to pair two items out of an answer collection set. True/false questions consist of statements that students have to decide whether they are true or false. The design of a multiple-choice question is that the question offers a number of choices. Only one of the choices is correct, and the rest of the choices are set as distractors.

Although the use of multiple-choice-based assessment are widely used in educational assessment, they are less efficient to measure student learning as essay-based assessments. However, many assessors prefer multiple-choice tests because they are suitable to measure student basic information and factual understanding [13].

2.1.2 Multiple-choice Questions

The growth of educational institutions increases the number of candidates performing for a test [14]. This made it more difficult for the assessors to measure students understanding because as the number of candidates increases in an assessment as it takes more time to release the assessment results. Therefore, this brought the need for a new method to assess students quickly and effectively [15]. In order to cope with this issue in educational assessments, multiple-choice question was introduced in 1950s. According to [16], the idea came first by Frederick J. Kelly at the University of Kansas in a favour to discover the US talented recruits of World War I in 1914 because it is an effective and quick method. Later, multiple-choice question was spread out to schools and universities in early 1950s. Recently, it has been estimated that more than half a billion multiple-choice tests is set per year in only North of America [16]. That number shows a massive demand on multiple-choice-based tests as one of the main methods to assess student learning.

Multiple-choice question offers to students several choices and students are required to select the most appropriate choice. A typical multiple-choice question consists of three fundamental elements:

1- Item stem: this element usually shows the problem in a question form.
2- Key option: the key option is referred as the correct answer.
3- Distractor: the distractor options is set to deceive students from the correct answer.

It has been argued that multiple-choice questions do not fully test student understanding as it seems sometimes very simple or it can be solved by guessing. Clegg and Cashin [8] claim that well-structured and designed multiple-choice question can measure high levels of student understanding that includes the six levels of Bloom's taxonomy of cognitive objectives. Bloom's taxonomy is designed to guide educators on how to assess students by setting several hierarchy levels as shown in Figure 2.1. Bloom ordered the levels from the bottom to top as follow: Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. Every level has to be mastered by the students before the next level can be attained. Therefore, assessors could design multiple-choice questions to test students on lower learning levels like knowledge or comprehension level, which makes some multiple-choice questions seem to be superficial for the students, or they could design multiple-choice questions-based on higher learning levels such as synthesis or evaluation, which makes multiple-choice questions seem to be more challenging for the students [17].

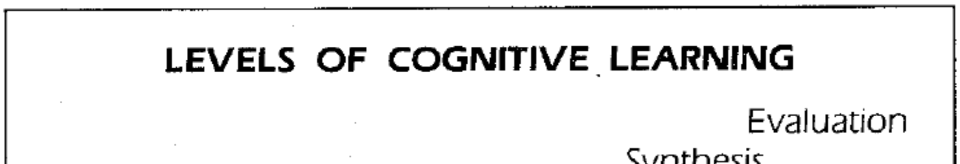


Figure 2.1: The six levels of Bloom's taxonomy of cognitive objectives [8].

Roberts [18] states in his research on the use of multiple-choice questions that the strengths of multiple-choice questions are:

- Multiple-choice questions can measure all the levels of the learning from the low to the top level.
- Multiple-choice questions can reveal the common mistakes if the incorrect choices in each question are well designed.
- As many multiple-choice questions can be used, they can be used to cover many subject materials.
- Multiple-choice questions require short time to answer when they are compared to any question type.
- One of the most advantageous point about multiple-choice questions is that an assessor or a machine can mark them accurately and easily.

Whereas Roberts carries out the main disadvantages of multiple-choice questions, which can be listed below:

- Multiple-choice tests take longer time to design compared to the rest of test types.
- Multiple-choice questions knowledge and do not test analysis skills and critical thinking of students.
- Multiple-choice questions push students to adapt simple approach to learn.

Roberts argues that the disadvantages, which are listed above, can be reduced easily if assessors focus more on implementing innovative multiple questions.

2.2 Automated Multiple-choice Scoring

Assessments have an important role in measuring student level of understanding. Scoring paper-based assessments is challenging as designing them. Scoring too many test papers for assessors is frustrating, time consuming, and not free of mismarks. Moreover, for students, they have to wait for long time to receive their feedback on how they performed and what their mistakes were. The problem arises when there are various parts of a test requires different markers, regardless the cost of the test markers, this implies that all the markers have to coordinate to reduce the marking time [19].

In an effort to reduce marking time, there has been significant works done in this field, which is called an automated scoring system. In this section, an overview of automated multiple-choice test machine and optical mark recognition technique will be explored.

2.2.1 Multiple-choice Test Machine

According to IBM [20], the first spark for an automated multiple-choice scoring machine was ignited in Ironwood, Michigan by Raymond Johnson in early 1930s. He was trying to grade his student test papers, as he was a high school physic teacher. He experimented a machine using small electrical circuit that detects a pencil mark and compares it with a key mark. Later, he ended up working in IBM as an engineer and became one of IBM's most innovative engineers. Johnson's work gave IBM the basis to create the first commercial test scoring machine that is called IBM 805 Test Scoring Machine in 1937 as shown in Figure 2.2. It was a breakthrough in educational advances as shown in in Figure 2.3 an image of the product advertisement.

Figure 2.2: A commercial image of IBM 805 Test Scoring Machine [20].

The process of the machine starts by inserting student answer sheets (cards) into the machine. The machine reads the pencil marks by sensing the pencil lead as lead makes connectivity when it exposed to electrical charges through a contacts plates. There is a “scoring key” determine the right and the wrong answers base on the connectivity of the shaded answer and then the machine shows the total score via an inductor. The machine needs a human intervention, as it cannot score multiple test cards automatically. It was estimated that the machine scoring rate is about 800 cards in an hour if an experienced operator operates it.

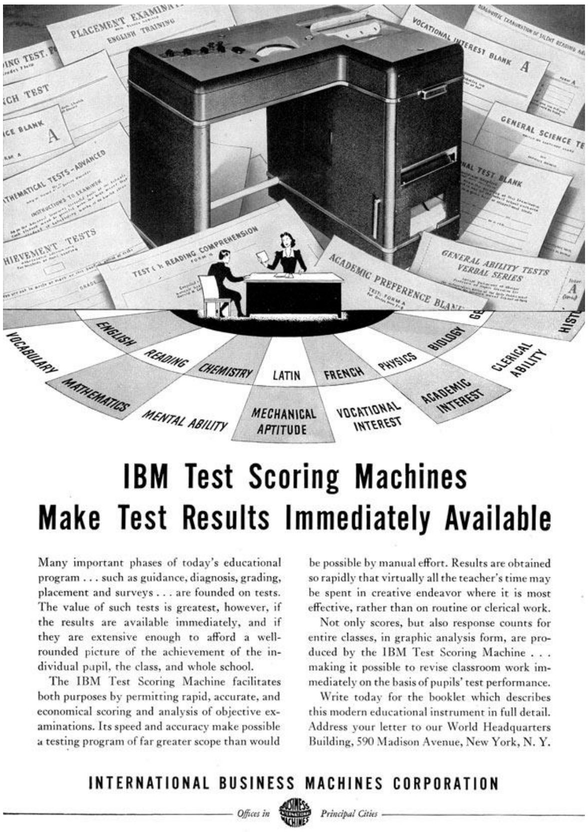


Figure 2.3: IBM 805 Advertisement [20].

Later in the beginning of 1960s, Johnson’s method was replaced by a new technology called optical mark recognition. The first optical mark recognition was designed successfully by professor Everett Franklin Lindquist. Lindquist’s mechanism is that the acquired marks are contrasted by a light beam at the mark positions of the answer sheet to recognise the selected choice. The machine recognises the shaded mark as it reflects less light compared the unshaded mark on the answer sheet [21]. The advantage of the optical mark recognition machine over IBM 805 was the speed and accuracy of the processing because Lindquist's machine utilised vacuum tube technology, which was measurably faster than a traditional electrical circuit. It was estimated that Lindquist's machine could mark 800 answer sheets per hour. Later, IBM designed another optical scoring machine called IBM 1230 as shown in Figure 2.4, which was based on Lindquist's technique. It paved the way for optical scoring machine and increased the popularity of scoring machines because it was more reliable and faster. It had the ability to score up to 1600 answer sheets. And not just like IBM 805, IBM 1230 does not require an operator to insert the answer sheets one by one. It had the ability to score multiple sheet automatically [22].

Figure 2.4: The first IBM optical mark recognition machine (IBM 1230) [23].

Later, many companies proposed optical mark recognition machines with the same concept to be used in different applications such as automated ballots and research surveys. According to Patel and Prajapati [21], a traditional optical mark recognition machine mainly consists of these three sections:

- Feeding section: The role of this section is to pick up the answer sheets and preper them to be proccessed for the next section (Photoelectric Conversion section). An example of a tradtional answer sheet is shown in Figure 2.5.
- Photoelectric Conversion section: The process of this stage starts by emiting light to the answer part of the sheet and the intensity reflection of the answer is transofromed into an electric signal via a light dependant sensor "photo sensor" as shown in Figure 2.6. Next, the elecrical signal is stored in an image memory.

The photosenser is set to have an intensity threshold to indicate ‘0’ for a bright light refelection and ‘1’ for a dim light reflection.

- Control section: The stored data in the image memory are processed by a recognition processor. This unit takes the recognition part based on the key answer.

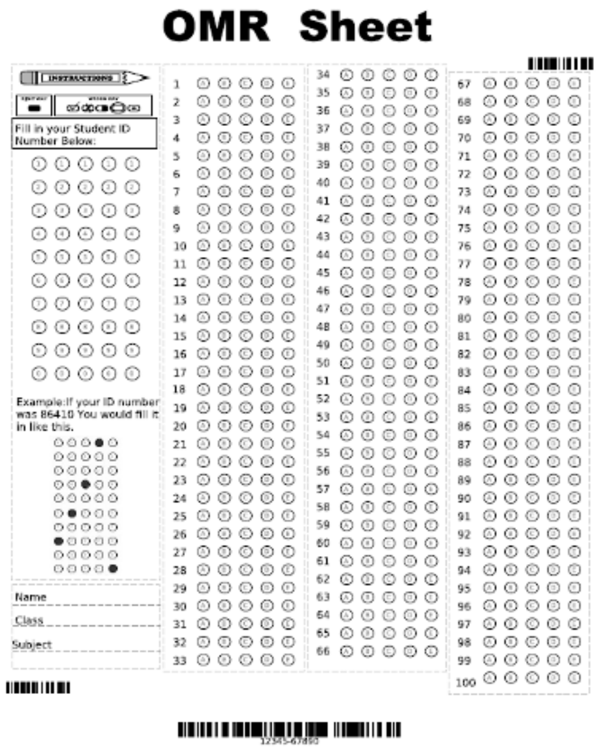


Figure 2.5: A traditional answer sheet of an optical mark recognition system [24].

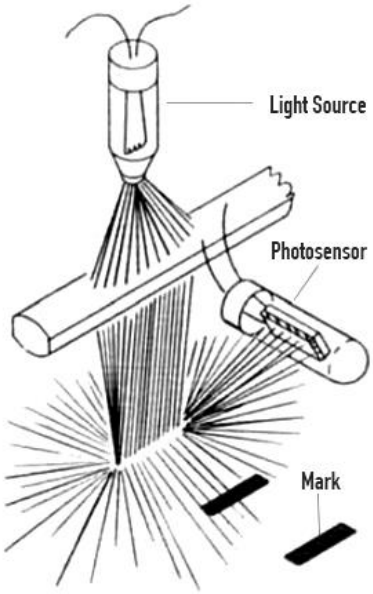


Figure 2.6: Photoelectric Conversion unit of IBM 1230, which contains a light source to emit light and a photo sensor to detect the reflected light [23].

2.2.2 State of the Art of Optical Mark Machine

In the last two decades, there are several organisations offer optical mark recognition solutions. In this sub-section, Scantron optical mark recognition products will be reviewed because Scantron Corporation is recent leading companies in optical marks recognition machines [25]. Although Scantron optical readers carry the same idea of IBM optical mark recognition machines concept, Scantron's optical mark recognition system was patented. What makes Scantron Corporation special is that Scantron Corporation focuses on providing low cost optical mark readers to educational institutions [25]. Scantron claims that it serves 98% of the best American schools and 94 universities of the best universities in America, and their products are utilised in different 56 countries across the globe. The growth in the demand of Scantron products shows that Scantron optical mark recognition solutions are reliable for the assessors [26]. Scantron categorises its scoring machines into three volumes:

- Low volume: The scoring machine can process 2800 to 3400 sheets per hour.
- Medium volume: The scoring machine can process 3600 to 6600 sheets per hour.
- High volume: The scoring machine can process 7000 to 15000 sheets per hour.

For the low volume series, Scantron offers three scoring series, which are EZData, iNSIGHT 20, and iNSIGHT 4ES. According to EZData brochure [27], EZData is designed for high accuracy results with a quick score feedback. Moreover, is design to be small and compact to make it more portable and light. The main feature of EZData is the low cost, which is approximately AU$150. The second scoring machine is iNSIGHT 20, which is designed to be one of the efficient and versatile because iNSIGHT 20 can be used as an optical mark recognition and image scanner. Not like EZData, iNSIGHT 20 has the capability to recognise accurately pencil and colour pen mark. Furthermore, iNSIGHT 20 has the ability to read Bar codes and hand written character such as student name and ID. According to the product brochure, the price of iNSIGHT 20 is higher than EZData, which is approximately AU$400 [28]. iNSIGHT 4ES capabilities are almost the same as iNSIGHT 20 but iNSIGHT 4ES series is designed to give the most precise results. Moreover, OpScan is capable to detect operation errors and fix skew in the sheets to give better answer registration. However, iNSIGHT 4ES is slower than iNSIGHT 20 as it can process 2800 sheets per hour. The price of iNSIGHT 4ES is AU$800 [29].

For the medium volume series, Scantron offers three scoring machine series, which are iNSIGHT 30, iNSIGHT 700c, and OpScan 8. The medium volume is generally designed to have higher speed compared to the low volume. iNSIGHT 30 is offered for institutions that look for higher marking speed at lower price. iNSIGHT 30 is able to scan and mark up to 4000 sheets per hour. Whereas, iNSIGHT 700c has the capability to process up to 6600 sheets per hour. Moreover, iNSIGHT 700c has the ability to detect and mark pencil and ink marks. These features put iNSIGHT 700c at higher price than iNSIGHT 30, which can cost up to AU$2000. OpScan 8 has almost the same iNSIGHT 30 features. However, OpScan 8 marking speed is less than iNSIGHT 30 as it can mark up to 5000 sheets per hour, which makes it cheaper than iNSIGHT 30 by a few 100 dollars [28].

There are three scoring machines in high volume line, which are iNSIGHT 70, iNSIGHT 150, and OpScan 16. All high volume machines are designed to work standalone without a need of a computer software to collect the scores. They are designed for higher capabilities to not just only mark multiple-choice tests, they can mark objective tests and surveys. The three offered machines from Scantron are almost the same. The differences are in the cost and the speed of the processing. The price range of the high volume machines is from AU$24000 to 30000 [30]. Figure 2.7 shows the images of all the volumes machines.

Figure 2.7: The complete set of Scantron optical mark scanners [26].

2.2.3 Software-based Automated Multiple-choice Test System

Unfortunately, because the traditional optical mark recognition machines are designed to be specialised scanners, the cost and setup time of the machines limited the use of the machines in only high enrolment units and in high budget institutions. Moreover, the traditional optical mark recognition machines require specialised forms (answer sheets) to mark. Most of the traditional optical mark recognition machines require a 'transoptic' paper as it allows the light to pass through it, which costs from US$0.15 to

0.2. Further restrictions, the traditional optical mark recognition machines restrict students to use a specific pen colour such as pencil. These restrictions pushed to open new possibilities due to the increment in learning environments.

In last the couple of decades, computer technology has evolved and dominated many fields of human life because it is cost effective and fast. Furthermore, image recognition methods have emerged such as pattern recognition, and image processing made computers smarter to make a decision. Recently, several new technologies offer solutions for document processing to recognise and process graphic images. The general concept of document processing technology can be summarised in Figure 2.8 [31].

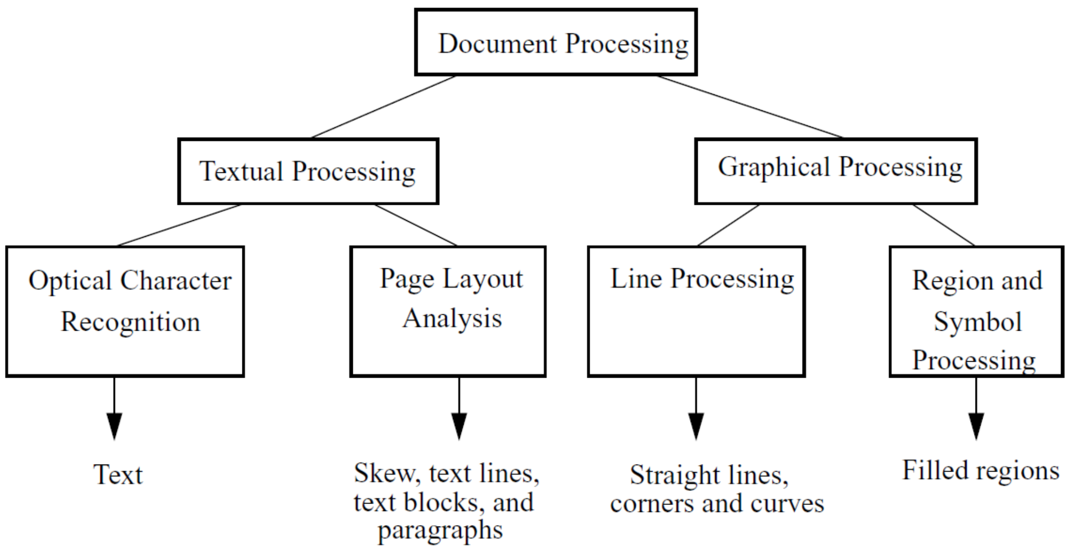


Figure 2.8: The main parts of document processing [31].

According to Sandhu, Singla, and Gupta [32], region symbol and optical character recognition were the base to achieve a new method of optical mark recognition system for an automated multiple-choice scoring system. The new solution aims for a cost effective and fast optical mark recognition system with customisable answer sheet. The reduction of the cost is because the new solution does not require a specialised optical scanner or sensors. It just requires an off-shelf scanner and a personal computer that is affordable to run the optical mark recognition software. The answer sheet is converted into a bitmap image by scanning it by an optical scanner, and the marks are recognised based on the aimed pixel intensity. This system of optical mark recognition can be called as a computer-based optical mark recognition. A computer-based optical mark recognition can be defined as a software operated via a computer that is designed to retrieve data from optical marks such as cross, ticks, and bubbles from a paper form [33]. As shown in Figure 2.9, the process of a computer-based optical mark recognition system starts by converting the answer sheet to a digital image via an ordinary optical scanner, and then the image is processed and marked via an optical mark recognition software, which is installed in a general purpose computer. Later, the processed data are stored in a database to be utilised later based on purpose of the user.

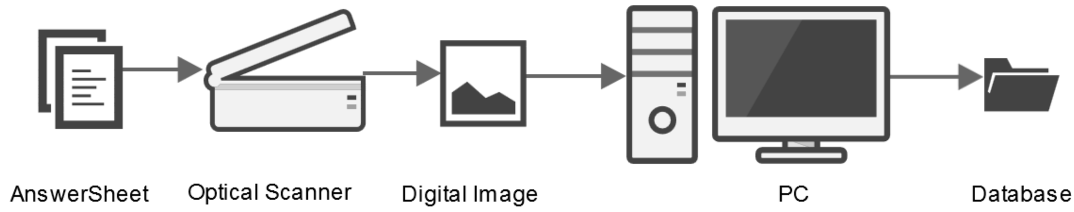


Figure 2.9: A computer-based optical mark recognition system, which consists of an ordinary scanner and a PC.

In addition, with the development of mobile phone platform, there are several mobile phone applications utilise mobile phone camera to perform an automatic marking same as the computer-based optical mark recognition. This system can be called as a mobile phone-based optical mark recognition system. Recently, there are several mobile phone applications run on Apple IOS and Android platforms such as OMR Evaluator on Android, Exam Reader on Apple IOS, and ZipGrade app on Apple IOS too.

The advantage of mobile phone-based over computer-based system is that the scanner, computer, and storage database are in one device. Moreover, mobile phone-based system is a mobile system which is more practical than a stationary system such as computer-based system. However, computer-based system is more reliable when there are many answer sheet to process because computer processors are more powerful than mobile phones. Moreover, nowadays, scanners can scan multiple papers at once, and it takes longer time to snap a large number of answer sheets. In addition, scanners have better quality than mobile phone cameras. Finally, mobile phone is not a reliable as a computer to score many papers. These mentioned reasons so far made computer-based system more common in educational institutions.

Overall, the main differences between optical mark recognition software (both computer-based and mobile phone-based systems) and optical mark recognition machines is the hardware in how they work. Optical mark recognition software works with common optical scanners or cameras as the main feature of software-based system. Whereas, optical mark recognition machines are designed to work with specialised tools such as pre-printed answer sheets and 2B pencil. Every optical machine works with a specialised pre-printed answer sheet to read it and in most of traditional machines, the answer sheet has to be marked by 2B type to recognise the marks. In contrast with the optical mark recognition software, the flexibility of creating an answer sheet is easier than the traditional machines. The answer sheet can be created and customised by any word processor. Moreover, the answer sheet can be later printed by any common printer and copied via any photocopier [33].

In a review of the related work of the optical mark recognition software, in 1999, Chinnasarn and Rangsanseri [34] developed the first PC-based marking system which reads printable answer sheets from an ordinary optical scanner. The downside of this technique is that an empty answer sheet has to be read first as learning model to recognise the interest areas such as unit and student code. Later, in operation model, a set of answer sheets are processed based on an answer model.

With the rise of digital and mobile cameras, Nguyen et al [35] developed a reliable algorithm to use camera instead of an optical scanner in aim to simplify multiple-choice questions marking. It is justified that acquiring answer sheets via a camera is faster and more portable than an optical scanner. Furthermore, size wise, camera is portable which is better for mobility when it is compared with a normal scanner size. However, the proposed system is only suitable for a small number of answer sheet. When scoring more than 100 answer sheets, using a camera to capture the sheets is tedious.

In Addition to Nguyen work, China, Assis, Neves, and Gonzalez [36] [37] developed a mobile phone that performs an automatic multiple-choice scoring on Android platform. The application was developed by OpenCV library. The scoring method is innovative which convert the answer sheet image into binary matrices. The application was tested in scoring hundreds of test papers.

In the same way, Karunanayake [38] developed an algorithm to score answer sheet which are acquired via web camera. Moreover, Karunanayake used template-matching technique to recognise answer sheets.

There is another Android mobile phone application which was designed by Tjahyadi, Budijono, Lukas, and Krisnadi [39]. The application design aims to provide an alternative solution. The results shows that the scoring rate is 100%.

Sanguansat [40] designed an algorithm that can recognise flexible answer sheet. The user is able to modify the answer sheet by using any spreadsheet software. The modification can be done to the number of answer items and the position of student name and ID matrix.

Čupić has implemented multiple solutions to score multiple-choice papers. First, Čupić [41] developed an open source Java-based marking system. The offline application is designed to mark two sets of test answer sheets; classical multiple-choice tests which offer multiple-choices without test questions in the sheets and integrated test and answers sheets which offer both but with less number of multiple-choice questions. The application offers different kind of information recognition for student ID and unit code such as barcode and matrix. Čupić et al [42] carried out with more emphasis on student ID identifier matrix in the same answer sheets to achieve a 100% recognition rate with the presence of high rotation and skewing. Furthermore, Bonačić et al [43] continued Čupić's research on information decoding for student ID by introducing an optical character recognition of seven–segment display digits. A template of a 7-segment display for each digit is designed for students to shade. The locations of digits and digit segments (areas of interest) are predefined to recognise numbers from 0 to 9 based on the input patterns. Although the method seems to be restrictive and based on shading instead of handwriting, it is more intuitive and easier than identifier matrix for student to encode as shown in results of over 90% success rate. Moreover, Čupić et al [44] devised a method for students to change their answers for a second or third time if answered wrongly. If students would like to alter their answer, they can annotate the error with a circle and write the correct letter next to the answer area. Later, while processing, if the error circle is shaded, the handwritten character is recognised whether it is A, B, C, E, or F character.

Chouvatut and Prathan [45] designed a multiple-choice algorithm. They put more emphasis on x-mark detection with different orientations. Their algorithm aims to be more practical and flexible with different pen colours.

Chai [46] designed an automated marking algorithm which emphasises more on result feedback. The proposed method marks printed answer sheets and annotates the scanned sheets by highlighting the correct and wrong answers next to each answer and send the annotated sheets back to student via email. The results show that the method is quick with up to 1.4 seconds per sheet and accurate. However, the proposed answer sheet does not enable student to change answer on the same answer sheet.

Bayar [47] [48] proposed an automatic grading scoring algorithm that depends on Hough transformation. He put more emphasis on registering marked answers. The algorithm was heavily tested by more than 1000 test paper to validate the algorithm robustness.

In a new approach to recognise unknown answer sheet form, Abdu and Mokji [49] designed an algorithm to detect and register marked answers from any type of answer sheets. In order to achieve this, an empty answer sheet has to be scanned first. Second, the answer sheet has to undergo pre-processing in aim to remove unwanted objects such as lines and text. Finally, a proposed combination of Region of Interest and Hough transform is used to determine the size and the locations of each answer box. The results show that the proposed algorithm is capable to recognise most common designs of answer sheets with high accuracy.

Afifi and Hussain [50] aimed to reduce the restriction even more on optical mark recognition. They put more emphasis on recognising answer box by using neural network. This offers the ability to deal with more shading styles. This technique is claimed to have higher accuracy and less restraints than most of the existing algorithms in optical mark recognition. However, this solution is not automated, and there is no such an emphasis on decoding student information.

Tang, Wang, and Xiao [51] used a new support vector machine method to recognise answer sheet. They put more effort to recognise multiple-choice answers in noisy environment and recognising user defined answer sheet design. It is claimed that their algorithm is more effective than the traditional answer sheet recognition.

Patel, Naik, and Ghare [52] design an optical mark recognition that is designed in VHDL for FPGA to recognise answer sheet. The aim was to design an alternative for image processing algorithms. They used IR sensors to scan answer sheets and process them directly via a VHDL script. This method is believed to be faster and cheaper than using a commercial scanner.

Zampirolli, Batista, Gonzalez [53] devised an automatic multiple-choice test generator and corrector. The aim of generator is to create a random multiple-choice questions which can be recognised from the key answer.

Haskins [54] implemented an image processing-based classifier for marked answer. The classifier goal is to distinguish selected answer from the deselected. Haskins conducted an accurate algorithm to detect the answers on the answer areas, and if there is shading outside the selected answer, the answer is presumed to be deselected.

Overall, from the aforementioned software-based scoring systems, it appears that there are several methods to implement a multiple-choice scoring system. Every scoring solution was designed to fit the purpose of the need. Table 2.1 summarises the ideas of the related works.

Table 2.1: A summary of image processing based contributions.

Authors	Year	Reference	Contributions
Chinnasarn and Rangsanseri	1999	[34]	The first PC image processing-based marking system.
Čupić and Bonačić	2011	[42][43]	Designed accurate ID identifier and introduced an optical character recognition of seven–segment display digits.
Nguyen	2011	[35]	A reliable algorithm to use camera instead of an optical scanner.
Abdu and Mokji	2012	[49]	Designed an algorithm to recognise unstructured answer sheet.
Tang, Wang, and Xiao	2012	[51]	A new support-vector machine method to recognise answer sheet.
Chouvatut and Prathan	2014	[45]	More emphasis on x-mark detection with different orientations.
Karunanayake	2015	[38]	An algorithm to score answer sheet acquired via web camera
Sanguansat	2015	[40]	An algorithm that can recognise partially unstructured answer sheet design.
China, Assis, Neves, and Gonzalez	2015	[36][37]	A mobile phone that performs an automatic multiple-choice scoring on Android platform.
Haskins	2015	[54]	an image processing-based classifier to recognise natural changing answer attempts
Chai	2016	[46]	More emphasis on result feedback via email
Patel, Naik, and Ghare	2016	[52]	Devised recognition algorithm via FPGA
Afifi and Hussain	2017	[50]	More emphasis on mark recognition via neural network.

2.2.4 State of the Art of Software-based Optical Mark System

One of the state of the art of software-based optical mark recognition can be seen in Remark Office OMR software for several reasons. Remark Office OMR software is the first scanning optical mark recognition software for gathering and processing marks from a plain answer sheet in the market with more 75000 clients across the world since 1991. Moreover, Remark Office OMR software can be said as the most flexible and capable for processing millions of answer sheets and forms [55]. The software was designed by Remark Products Group of Gravic. Gravic Group is an international company for collecting and analysing data solutions based in Pennsylvania, US. Gravic designed the software to be the complete solution to process any type of forms such as surveys, evaluation forms, tests, and ballot forms. The software is designed to have multiple standard forms. The forms are fully customisable to be modified to user purposes. The software is designed to serve:

- Education professionals for assessing students
- Marketing professionals for attaining customer's feedback
- Government departments for general surveying purposes
- Computer professionals for integrating with different software
- Human resource departments for employment applications
- And many more of special applications.

One of the main features of Remark Office OMR is that it is easy and flexible to use for a normal user without a need for additional experiences as any Scantron optical mark recognition machine. Gravic designed an easy system that consists of three steps: design, scan, and analyse. In the first step, the user has to design a form based on the targeted purpose. The form can be made by any word processing software such as Microsoft Word, Excel, and Adobe Photoshop or it can be made by using one of pre-made forms temples in the software. Later, the user can easily highlight question area via the software to tell the software where to recognise the marks. In the next step, the user can print and scan the form as much as wanted to be filled up based on the targeted purpose. In the last step, all the paper can be scanned via a normal optical scanner or via a multiple function scanner as Gravic recommends that for faster scanning. Later, all the scanned paper can be exported to the software to collect and analyse data based on the designed form. The system can be summarised in Figure 2.10. Figure 2.11 exposes the main window of the software, which show the database section, data manger, and scanned sheet preview field [56].

According to Remark Office OMR user manual [57], the software works on Microsoft Windows platform. The minimum requirements to run the software are a PC with at least 1 GHz or faster PC processer and at least 250 MB hard drive free space and 512 MB RAM are required to install the software. However, for better performance, Gravic recommends 1 GB free storage space and a supported printer from Windows for fewer issues.

The main feature of Remark Office OMR is to detect marks areas. However, there are further features of Remark Office OMR, which are:

- Built in optical character recognition feature to read machine printed texts
- Barcode recognition to read several barcode systems
- Built in analysis to provide statistics of the processed reports
- Data exportation to many document formats such as Dropbox, MS Excel, Google Dive, and SPSS.

With all these features in one software, the cost of Remark Office OMR is AU$ 995 for one time per user [56]. By comparison with Scantron optical mark recognition machines, Remark Office OMR seems to be a practical solution to mark many assessment answer sheets with a reasonable price as it can handle processing thousands of papers.

Figure 2.10: Remark Office OMR System, which consists of three steps: Design, Scan, and Analyse [55].

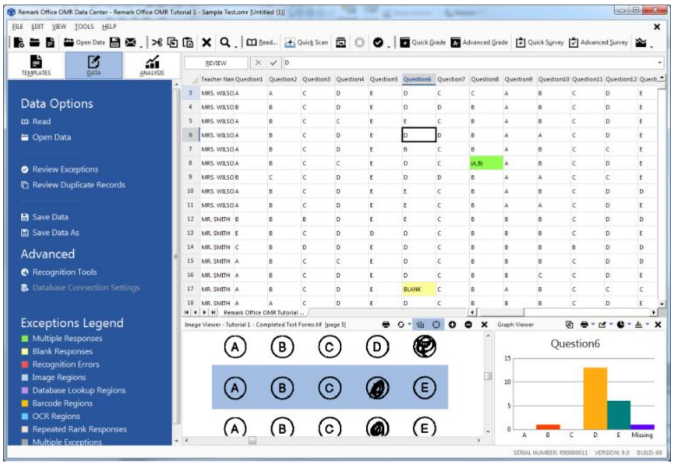


Figure 2.11: The main window of Remark Office OMR Software [56].

There is another software-based optical mark recognition that is worth mentioning as a state of the art in optical mark recognition, which is Eyegrade. Eyegrade has several features. The preponderant one is that Eyegrade system utilises a webcam instead of an optical scanner. Eyegrade take advantage pf a webcam to score multiple choice questions. As it just needs an off-shelf webcam, its main objective is to be a low-cost and portable solution. Eyegrade developers targets small scale educational institutes, and it can be handy to score a small number of students from 10 to 40.

There are several others features of the systems which are:

- **Scoring tests:** By obtaining a webcam, the graphical user interface of Eyegrade allows you to score tests. Eyegrade is designed to recognize not only the multiple choice questions, but also student IDs by using its hand-written digit recognition structure. Moreover, the graphical user interface allows the user to intervene if any scoring error happens as shown in Figure 2.12.
- **Exporting scores:** scored test papers marks can be exported in CSV format, which is widely recognisable, by many software such as MS Excel.
- **Test paper generation:** Eyegrade offers the ability to creating MCQ test or let the user create the test paper other software. It is able to create your MCQ test in PDF format. Furthermore, it can automatically build multiple versions of the test by shuffling questions and the choices within the questions

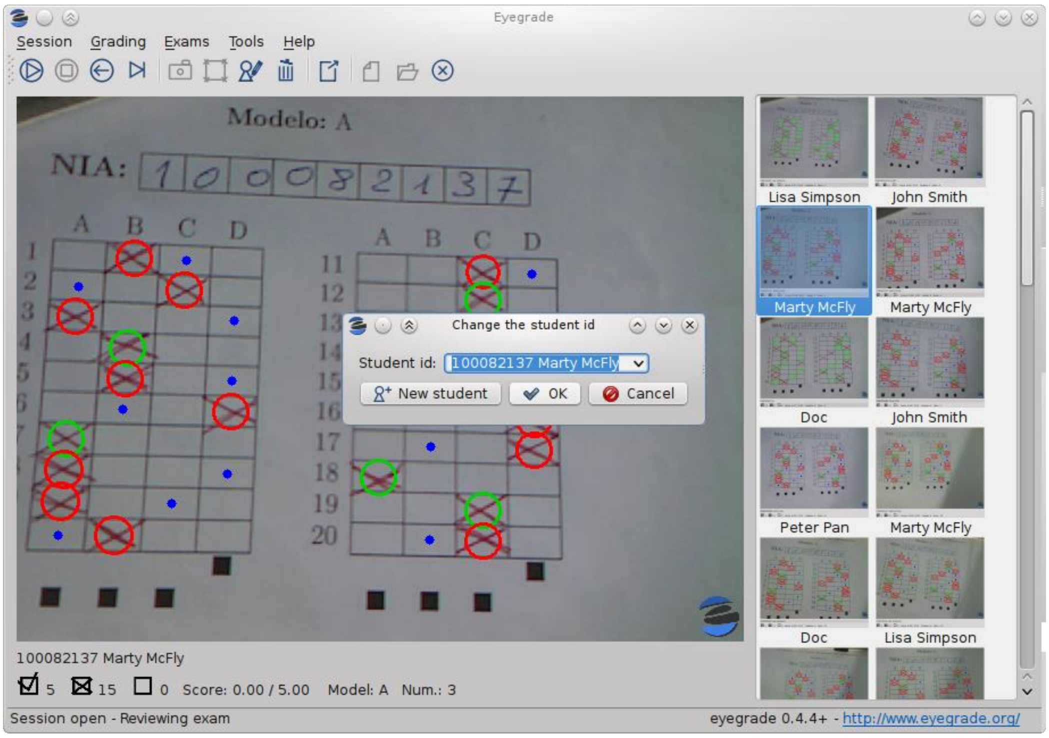


Figure 2.12: the graphical user interface of Eyegrade.

A demonstration video of the system flow can be view on the link (https://youtu.be/ajK9JAhsDH0). It is worth mentioning that Eyegrade is a free software and is licensed. Furthermore, it is an open source application, which means any developer can integrate their software features to theirs [7].

When it comes to comparing the both systems (Remark Office and Eyegrade), they are both useful based on the level of the use. For large-scale institutes such as universities, Remark Office is definitely the better solutions to fit, as it is design for a large number of exam papers. In the other hand, for a small-scale such as a small class test or a quiz, Eyegrade is the winner.

2.3 Image Processing Fundamentals

As the research aims to implement an image processing-based scoring system, image basics and document image processing will be explored.

2.3.1 Image Basics

Pixel Structure

After scanning an answer sheet, the sheet is converted into a digital image. A digital image consists of a set of pixels in 2D plane. Every single pixel in an image can be represented by denoting two values as presented in Figure 2.12. These two values represent an x position and a y position. The x and y values can be represented in 2D vector as shown in (2.1). For pixel coordinate, the origin is in the top left of the coordinate. The x value increases downward and the y value increases to the right [58]. The general notation of an image can be defined as

$$\mathrm{I}(\mathrm{x}, \mathrm{y}). \tag{2.1}$$

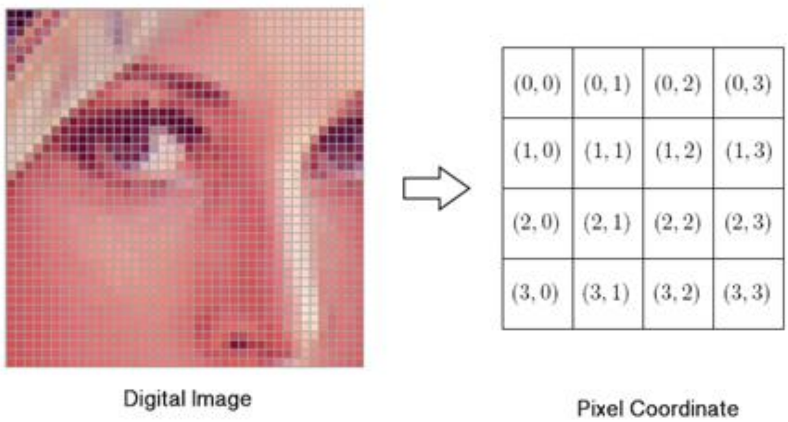


Figure 2.12: An example of a digital image and its pixel coordinate [59].

Each pixel has intensity value to represent a pixel in an image. There are two types of digital image: grayscale image and coloured image.

Image Colours

A scanned answer sheet could be scanned to be a colour image or a grayscale image. Hence, the difference between them will be explained. Grayscale image has different intensities in each pixel from black to white (grey level) as shown in Figure 2.13. The grey level shows the entire range from black to white. An individual pixel of eight-bit image holds an intensity value of 256 grey scales [60].

Figure 2.13: Pixel grey level of black and white image [60].

Each individual pixel of a coloured image holds three intensity values that represent red, green, and blue colour (RGB). Each colour has 256 intensity levels which means every colour hold 8 bits as shown in Figure 2.14. Therefore, each coloured pixel contains 24 bits of memory. RGB colours can be mixed in each pixel to create any possible colour such as yellow, cyan, or magenta [60].

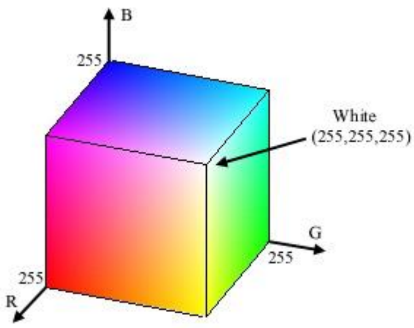


Figure 2.14: RGB colour space [61].

Image Resolution

Image resolution refers to the amount of pixel density in an image. If an image has high resolution, the image has higher sharpness, more details, and bigger size. An example is shown in Figure 2.15 demonstrates the differences when an image resolution is increased in term of details and sharpness [62].

Figure 2.15: An example of an image with different resolutions [60].

Image Transformation

Processing answer sheet involve transforming the image such as scaling, rotating, and translating. Image transformation is applied to map an image coordinate to another image coordinate system. Image transformation is useful when multiple images are captured from different positions. In addition, it is a powerful method to correct image distortions. A basic image transformation equation can be written as

$$\acute{x} = T\boldsymbol{x}, \tag{2.2}$$

where T is a transformation matrix $(n \times m)$, x is a pixel position in a space, and $\acute{x}$ is the pixel position after transformation which is called a spatial transformation. In addition, as shown in (2.3), it can be rewritten in a matrix form

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \tag{2.3}$$

where $a, b, c, and\ d$ are transformation parameters. The matrix (2.4) can be rewritten as an equation for $\hat{x}$ and $\hat{y}$

$$\hat{x} = ax + by,\ \ \hat{y} = cx + dy \text{ [63].} \tag{2.4}$$

There are two main types of 2D image transformations, linear transformation and nonlinear transformation. If a linear transformation is applied to an image, the image retains its diminutions. Whereas, the nonlinear transformation changes the image diminutions. The scaling factor determines whether a transformation is linear or nonlinear. A transformation can be linear if the value of x and y scaling factor are equal. If one of the parameters is different, the transformation is called nonlinear as shown in Figure 2.16. The scaling factor can be represented as

$$\grave{x} = S\,x, \tag{2.5}$$

$$\begin{bmatrix} \grave{x} \\ \grave{y} \end{bmatrix} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \tag{2.6}$$

where s_x is a scaling factor for x and s_y is a scaling factor for y [64].

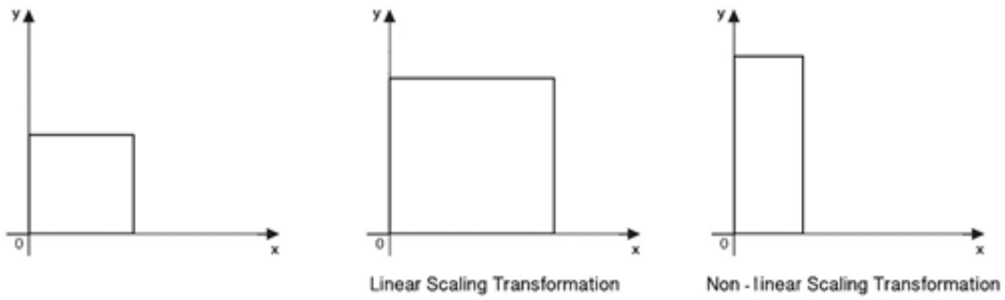


Figure 2.16: An example of linear scaling and non-linear scaling.

The basic 2D image transformations are:

- Translation
- Rotation
- Shear.

2D translation transformation can be represented by 2 x 3 matrix and an example of simple translation is shown in Figure 2.17. 2D translation can estimated as

$$\grave{x} = [I \;\; t]\, \boldsymbol{x}, \tag{2.7}$$

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tag{2.8}$$

where t_x is a translation for x and t_y is a translation for y [63].

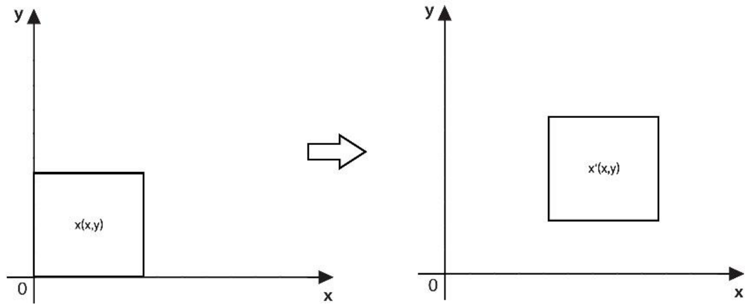


Figure 2.17: An example of simple 2D translation.

For simple image rotation about the origin as in Figure 2.18, the rotation can be represented as

$$\grave{x} = R\,\boldsymbol{x}, \tag{2.9}$$

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \text{ [63]} \tag{2.10}$$

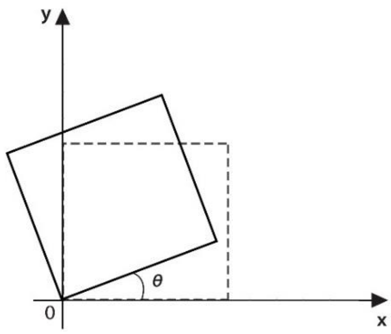


Figure 2.18: An example of a rotation about the origin.

In shear transformation, one axis is moved downward or upward to create parallel motion as shown in Figure 2.19. If the shear is toward x-axis, the value of $S_y = 0$. Moreover, if the shear is toward y-axis, the value of $S_x = 0$ as shown in Figure 13 [63]. Shear transformation can represented as

$$\grave{x} = S\boldsymbol{x}, \tag{2.11}$$

where S is 2 x 2 matrix as shown (2.11),

$$S = \begin{bmatrix} 1 & S_y \\ S_x & 1 \end{bmatrix}. \tag{2.12}$$

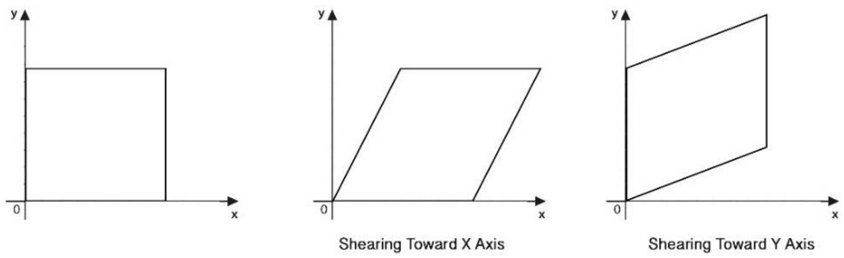


Figure 2.19: an example of a shear in X and Y-axis.

Image transformation can be more complex, which involves translation, scaling, and rotation in a same motion. The image aspect ratio will be different when the transformation is applied which can be defined as

$$\grave{x} = sRx + t \text{ or } \grave{x} = [sR \;\; t]\, x, \tag{2.13}$$

$$\begin{bmatrix} \grave{x} \\ \grave{y} \end{bmatrix} = \begin{bmatrix} s_x \cos\theta & -s_y \sin\theta & t_x \\ s_x \sin\theta & s_y \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \tag{2.14}$$

This kind of transformation is called affine transformation [64], which can be written as $\grave{x} = Ax$. It can be represented by a 2x3 matrix as

$$\grave{x} = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \end{bmatrix} x. \tag{2.15}$$

2.3.2 Document Image Processing

Image processing is a wide topic. As our research is more concerns about document analysis, in this part of the sub-heading, several concepts about document image processing will be analysed.

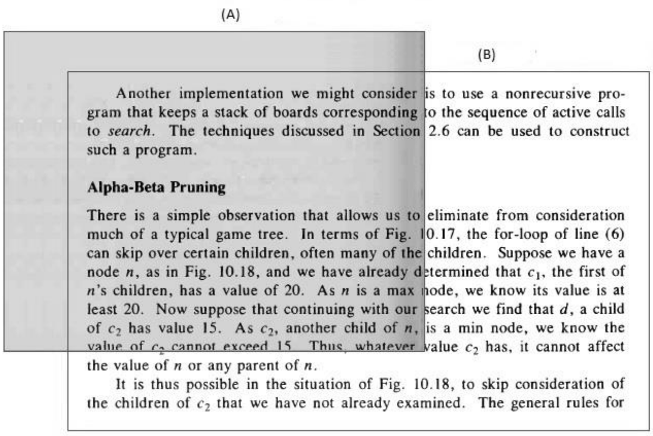
Image Binarisation

Image binarisation is the main processing step for answer sheet processing and recognition. Image binarisation is defined as the process of converting grayscale image pixels into black and white [65]. Every grayscale image pixel from 0 to 255 is set to 0 (white) or 1 (black) based on a predetermined threshold. Binarising threshold can be represented as

$$g(x,y) = \begin{cases} 1, & I(x,y) > \mathrm{T} \\ 0, & I(x,y) < \mathrm{T} \end{cases}, \tag{2.16}$$

where $g(x,y)$ is the binarised pixel, $I(x,y)$ is the grayscale pixel, and T is the binarisation threshold.

As documents have a single background, the aim of paper binarisation is to separate or segment answer sheet contents from the background for further processing. An example of a document image binarisation is shown in Figure 2.20 [66].



Another implementation we might consider is to use a nonrecursive program that keeps a stack of boards corresponding to the sequence of active calls to *search*. The techniques discussed in Section 2.6 can be used to construct such a program.

Alpha-Beta Pruning

There is a simple observation that allows us to eliminate from consideration much of a typical game tree. In terms of Fig. 10.17, the for-loop of line (6) can skip over certain children, often many of the children. Suppose we have a node n, as in Fig. 10.18, and we have already determined that c_1, the first of n's children, has a value of 20. As n is a max node, we know its value is at least 20. Now suppose that continuing with our search we find that d, a child of c_2 has value 15. As c_2, another child of n, is a min node, we know the value of c_2 cannot exceed 15. Thus, whatever value c_2 has, it cannot affect the value of n or any parent of n.

It is thus possible in the situation of Fig. 10.18, to skip consideration of the children of c_2 that we have not already examined. The general rules for

Figure 2.20: Document image binarisation example, where (A) is the grayscale image and (B) is the binarised image [67].

If the background pixels are consistent throughout all the image, using one threshold is enough. The single threshold is called a global threshold. In many cases, a global threshold is the best to segment the background from the document contents. For instance, if a document contains a simple white background as shown Figure 2.20, applying a global threshold is adequate. Global threshold is the simplest way to binarise an image. Based on the required threshold, all the image pixels will be binarised. However, global threshold has several disadvantages. First, if a document has a poor contrast between the background and the contents, the global threshold fails to segment the content. Moreover, if a document contains some small contents such as small texts, they will be eroded after applying the threshold. Therefore, in these cases, a smarter threshold is needed.

An effective method to solve different intensity issue is to apply an adaptive threshold. Adaptive threshold analyses image pixels intensities and calculates the average threshold. Basically, the threshold changes with the change in pixel intensity. An example of an adaptive thresholding is shown in Figure 2.21 [68]. There are various algorithms and techniques on how to calculate the thresholds. One of the common algorithms is that estimating the white and black scales and then subtract them from the original image. Hence, the image is left with brightest and darkest pixels. Figure 2.22 illustrate the adaptive threshold algorithm [31].

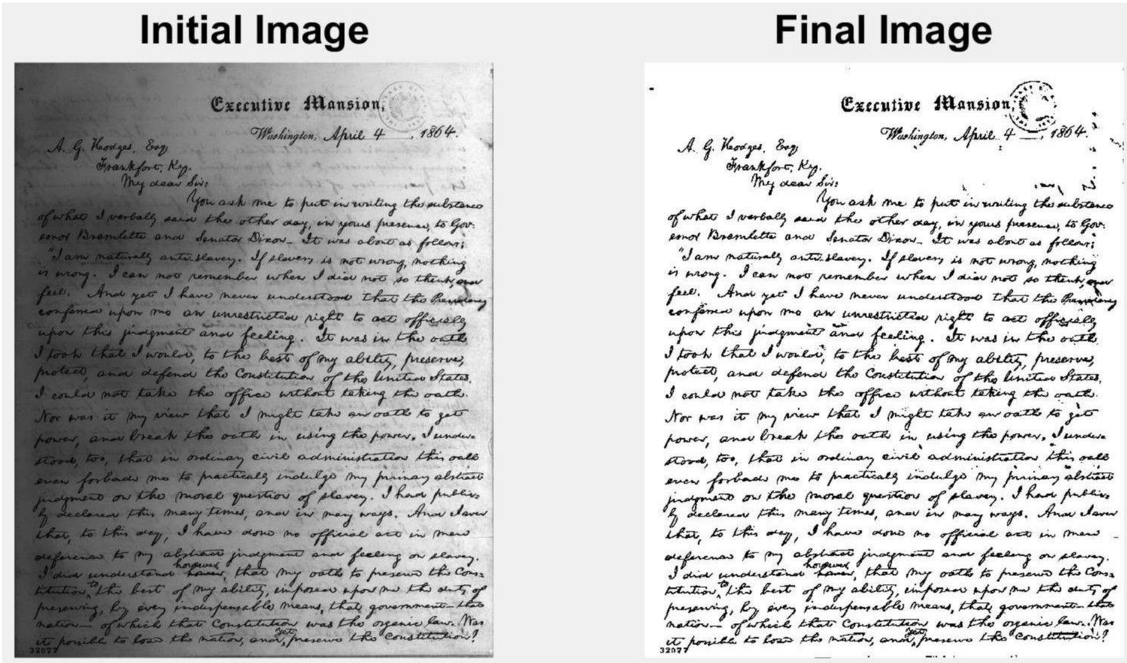


Figure 2.21: An example of applying an adaptive threshold [67].

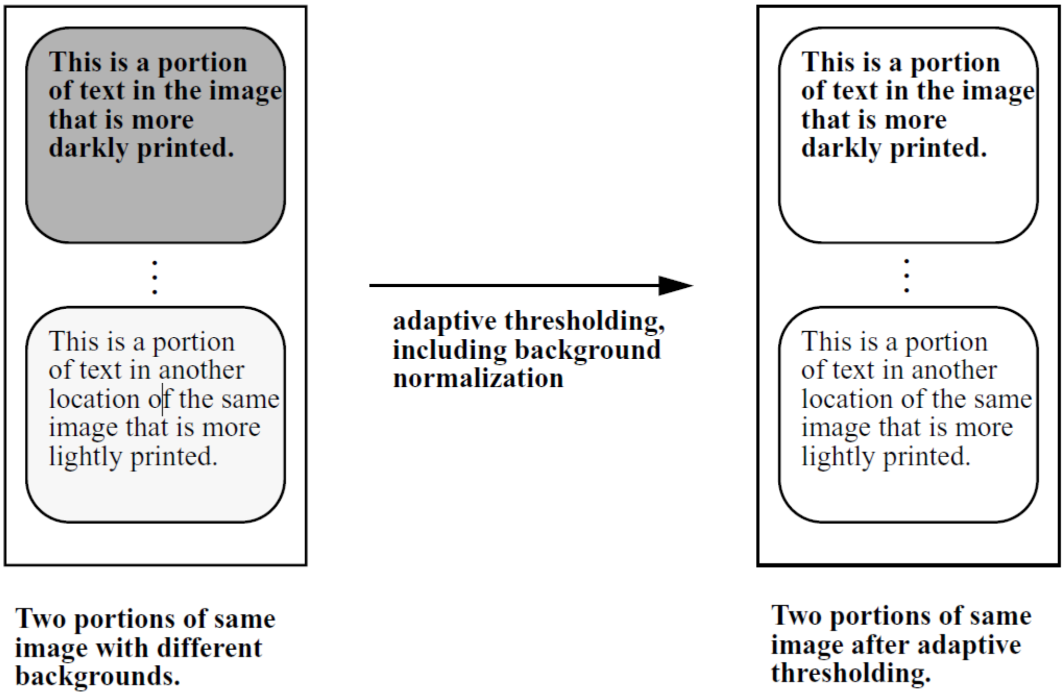


Figure 2.22: An adaptive threshold algorithm with two different intensities [31].

Choosing a threshold method whether it is global or adaptive threshold, is all based on the application of use. It depends on the quality of the document image and how it was acquired via a camera or a scanner. However, the processing of binarsing is now usually perfect. Noises after applying a threshold is always expected whatever threshold is used. Moreover, content deformations, such as disconnected thickened text, may occur depending on the value of the threshold. It is recommended to apply a noise removal and thinning operation to enhance the context.

Noise Removal

Noise is one of the main issues faced after binarising a document image [69]. The most common types of noises are salt and pepper noise and hole noise as a result of binarising. An example of salt and pepper noise and hole noise is shown in Figure 2.23. The noise usually increases due to poor binarisation. The noise occurs in the isolated area as shown in the figure. In addition, the noise could happen in the interest area of the context as holes. The main reason to reduce the noise is that a failure to erode the noise will lead to a faulty recognition result or a wrong perhaps. In every noise removal algorithm, the aim is to reduce the unwanted noise as much as possible and retain the original contents [70].

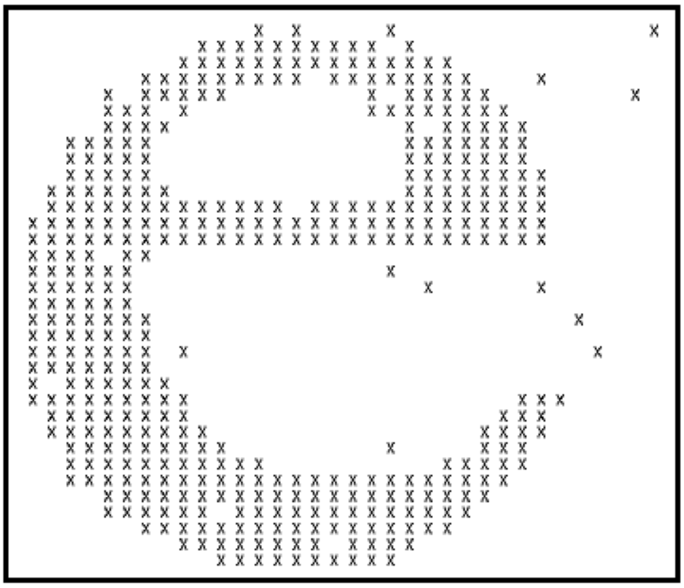


Figure 2.23: Letter "e" exposed to pepper and salt noise, where x represents ON pixels [31].

There are many noise removal algorithms, and every algorithm fits its purpose. However, there is a generic method. The generic method based on pixel erosion and dilation operations. Pixel erosion operation is defined as the process of shrinking a group of pixel boundary based on a threshold which determine the size of the pixel boundary. And pixel dilatation is defined as the process of enlarging a group of pixel boundary based on a threshold which determine the size of the pixel boundary.

The generic method simply works by erode a pixel group first to eliminate the salt and pepper noise in the isolated area of a document image. Later, the pixel groups are dilated back to retain the shrinked pixel groups. Or, process can be reversed to fill up the holes in the contents by first dilating the pixels to fill the holes and then eroding the pixels. The both ways techniques are illustrated in Figure 2.24. As in the figure, (A) is a pixel group with a salt and pepper noise which is removed by erosion-dilating method. In addition, (B) is an opposite noise removal to fill the hole of a pixel group.

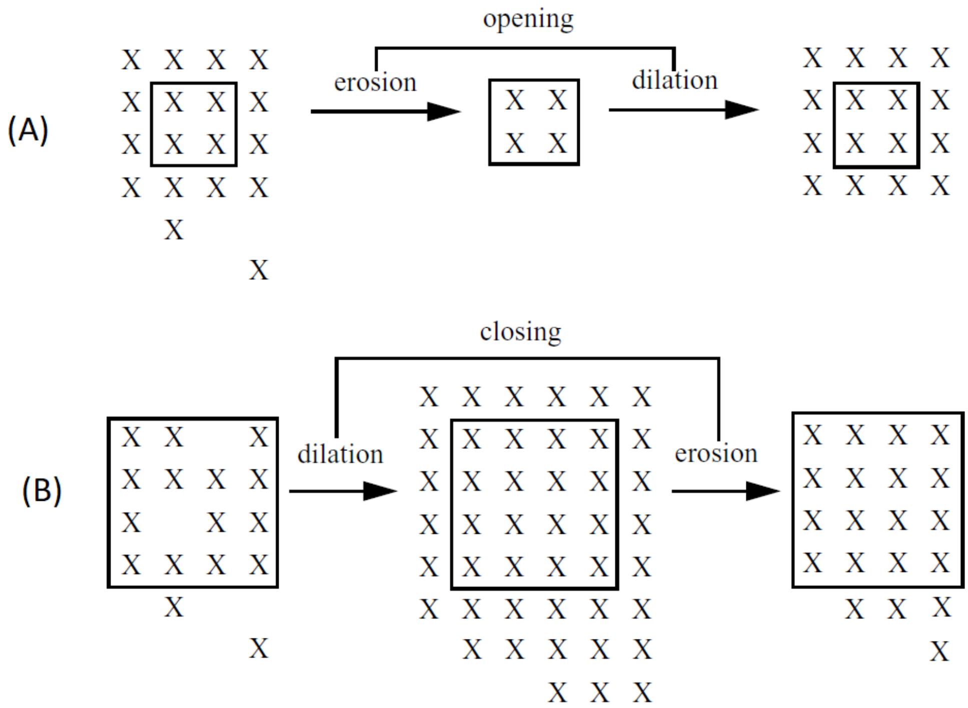


Figure 2.24: An example of a noise removal with salt and pepper noise and holes noise [31].

Thinning and Chain Coding

Usually after applying a noise removal, thinning operation is used as part of document image analysis. Thinning operation can be defined as the process of reducing the lines of a binary region to retain the skeleton of the region. The aim of the operation is to retrieve the main part of a region in an image, hence further processing can be made to extract the features of the region. Thinning operation may be called in different references as “skeletonising”. Thinning operation is commonly part of pattern recognition algorithms. For example, if a word is written with different pen thicknesses, thinning helps to retrieve the same pixel regions. Although there are several methods to track lines, thinning operation is faster to track as the regions are thinner. In most document image algorithm, thinning is part of in the pre-processing level of document image algorithm such as optical character recognition, electrical circuit detection, and fingerprint recognition. Examples of image thinning operation are shown in Figure 2.25 [71].

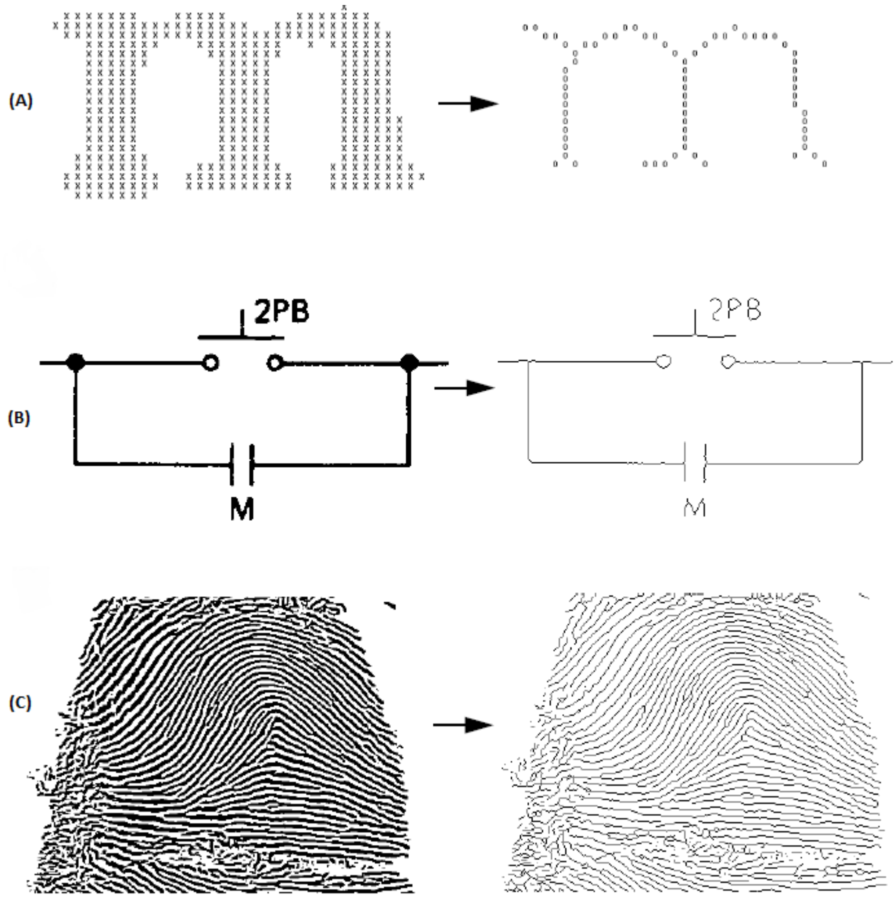


Figure 2.25: Examples of binary regions thinning, where on the left side is the binary region, and on the right side is the thinned region [31].

There are several requirements of thinning operation to achieve a fine skeleton of a binary region. Firstly, the binary region must be connected. Small disconnected region after thinning will be eroded. The binary region should have at least eight connected pixels to achieve a good thinning result. Furthermore, the result of thinning should create lines which are similar to central lines of the binary image. After applying thinning to small branches image such as a fingerprint (Figure 2.25 (C)), noise increases and should be reduced.

When a binary region is thinned, it can be easily determined the sequences of the binary region. This is achieved by labeling the connected neighbors in the 3x3 kernel of a line or a curve. This operation is called chain coding. Its benefit is that the direction of a line is labeled instead of labeling the whole binary region. Storage wise, chain coding is efficient to save the memory. More importantly, this operation can ease the processing for further operation such as line detection, edge smoothing, and feature detection.

There are many chain coding techniques. The most widely used is Freeman chain code. It starts first by searching from the top left until bottom right of a binary image. When it locates a binary pixel, the location of the pixel is stored. After running through all the image pixels, the sequence code of the searched pixels is stored from the first pixel location until the last neighboring pixel of the region skeleton [72].

3. METHODOLOGY

In the third chapter, the approach of the proposed system will be explained. This starts by the proposed system design in Section 3.1. Later, Section 3.2 will discuss the answer sheet design for the system. Moreover, the details of the scoring algorithm are discussed in Section 3.3. Finally, Section 3.4 will address the materials needed to implement the system.

3.1 System Design

The proposed system is set to meet the research objectives. The intended system is a computer-based scoring system. As discussed in the background about the software-based, the designed system requires the same equipment to utilise as illustrated in Figure 2.9. The system is designed as shown in Figure 3.1. The figure shows the high-level design of the system. As illustrated in Figure 3.1, the system consists of three inputs and three outputs. First, the system requires a list of student information, which are student names, IDs, and emails of the intended class. After collecting the answer sheets of a test, the filled answer sheets have to be scanned and then fed into the software algorithm to be scored. The proposed system adopts an optical scanner as opposed to camera. Despite the fact that the camera is better in mobility, the reason for choosing an optical scanner is that modern scanners can be used to scan a large number of answer sheets with a better quality and less noise. Furthermore, the recent available scanners are considered affordable especially for a large institute such as a university. The details of the chosen scanner will be discussed in Section 3.4. Lastly, the algorithm requires answer key as a reference to score. The imported student database and answer key files have to be in an Excel sheet format. In addition, the scanned answer sheet format must be in a JPG format.

After providing three inputs, the software algorithm will process the data. After executing the algorithm, the algorithm should provide three output as shown in Figure 3.1. The first output is an excel sheet of all students IDs with the corresponding test scores. The second output is annotated answer sheets, which contain highlights of the right and wrong answers. And lastly, the third output is emails to test takers which contain the total score and the annotated sheet.

There are some assumptions of the proposed system. First, all the answer sheets are scanned in A4 size and are not deformed. Moreover, all the three input of the system are provided with the right format. Lastly, the answer key and student database are right and free of mistakes.

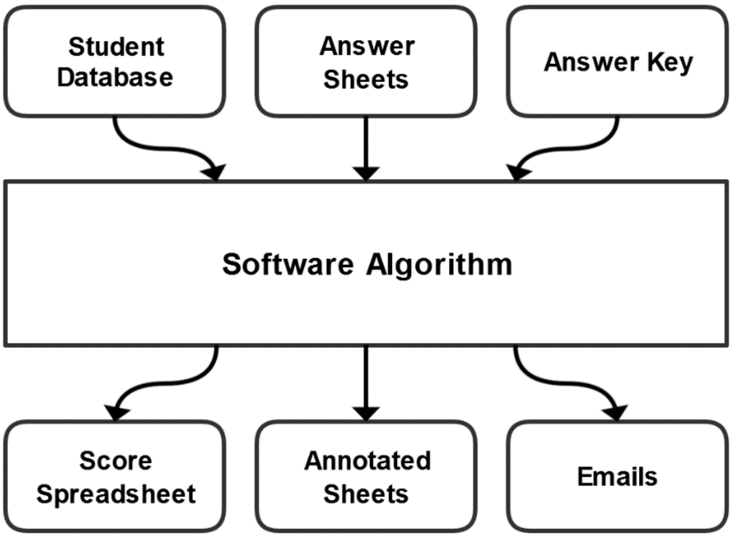


Figure 3.1: The flow of the proposed system.

3.2 Answer Sheet Design

Before going through the software algorithm details, the answer sheet has to be designed to meet the requirements of the software algorithm. The layout of the answer sheet design is shown Figure 3.2. From Figure 3.2, it can be seen that the proposed answer sheet consists of four sections which are Exam information, student ID & name, instructions, and answer area. Exam information area is an editable area where the assessor can add the exam date, unit title & code, and campus name before an exam via a PDF editor such as Adobe Acrobat. The instructions area has some guidelines for students on how to shade and how to write student ID & name in the designated area. The area where students can shade their answers is called answer area. Based on the mentioned guideline, the actual answer sheet was designed as shown in Figure 3.3. The answer sheet is designed to handle up to 72 multiple-choice answers, each item contains five choices. The answer sheet is designed to not only allow students to shade their answers, but also give the ability to re-answer instead of using an eraser. An example of how to change an answer on the answer sheet is shown is in Figure 3.3 in the instruction section.

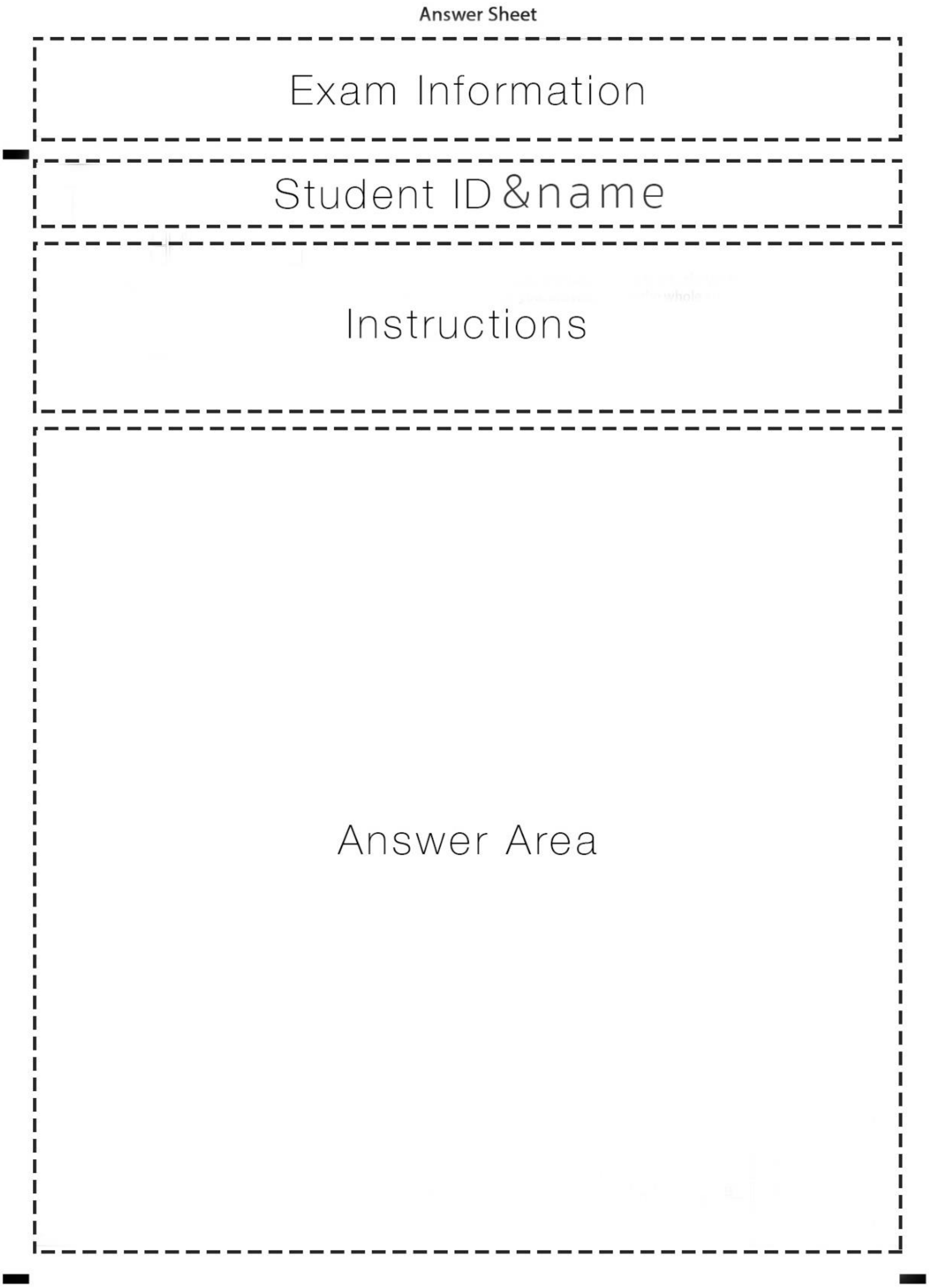


Figure 3.2: Answer sheet areas.

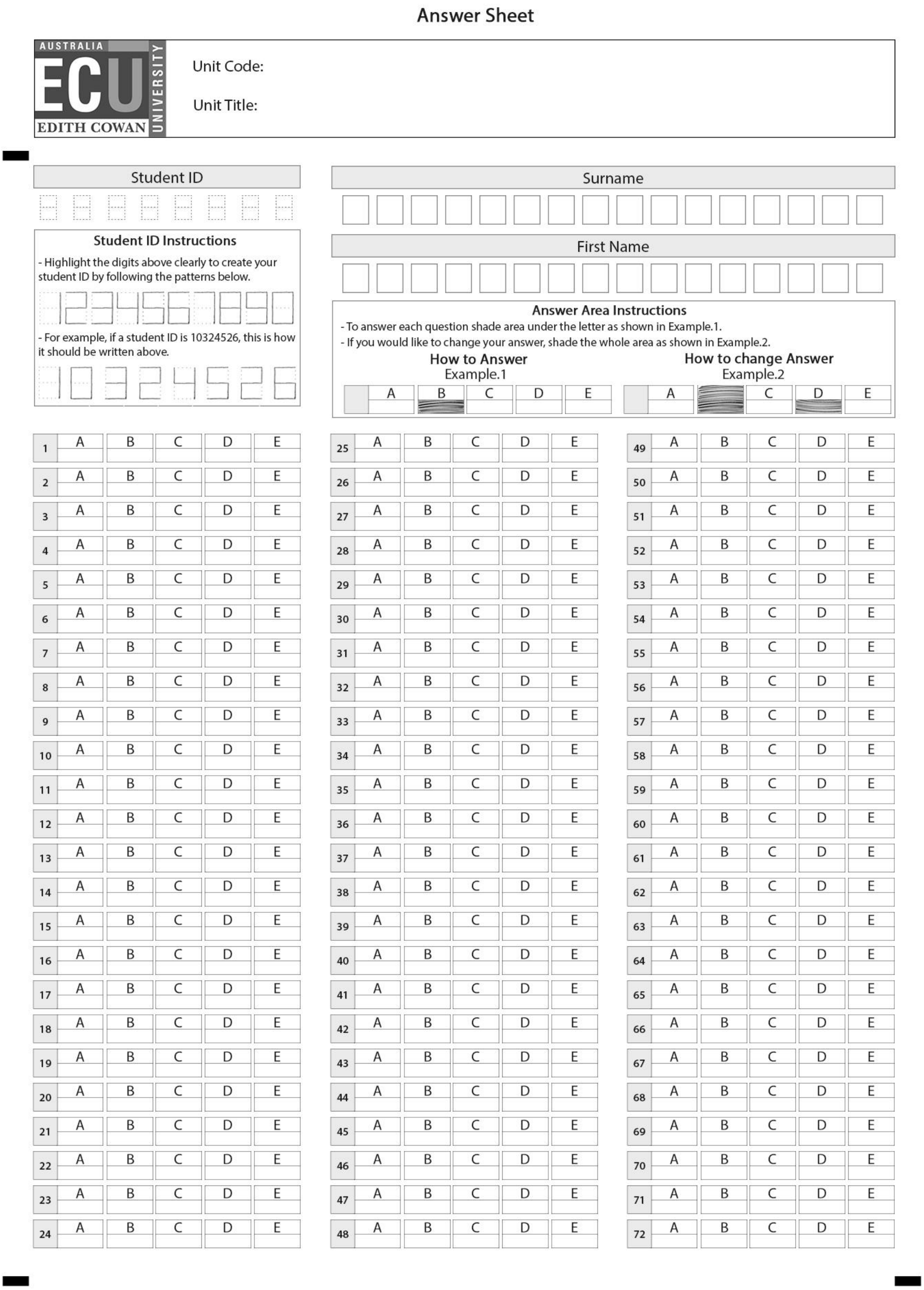


Figure 3.3: The proposed answer sheet design.

3.3 Software Algorithm Design

The software algorithm is the heart of the proposed system as it takes the bigger portion of the scoring process. The algorithm process is automated until the last input answer sheet as it is aimed in this research. The software algorithm is composed of four main algorithms which are interconnected to work together. The four algorithm are:

- Finder pattern algorithm
- Student information recognition algorithm
- Mark recognition algorithm
- Score feedback algorithm

The details of these algorithms will be explained in the following sub-sections. Figure 3.4 shows the software algorithm structure in a flow chart. The flow chart shows the overall work flow of the software algorithm. As shown in Figure 3.4, after providing the three inputs to the software algorithm, the algorithm starts by registering the answer sheet. The aim of the registration is to find the three finder pattern points. If they are registered, a tilt adjustment is made to straighten the answer sheet. After adjusting the answer sheet, student information area and answer area are segmented to recognise whose paper is and what student's score is. After segmentation as shown in Figure 3.4, the proposed method will execute a mark recognition algorithm to detect marks in answer area, as well as student information recognition algorithm to recognise student identity from the database. After decoding student information and scoring answer area based on the key answer correctly, the scored answer sheet will be annotated to show the correct and wrong answers with total score. Furthermore, a spreadsheet will be generated which lists each recognised student information with its corresponding score.

As shown in Figure 3.4, the same process will be repeated until the last scanned answer sheet in the specified directory. Finally, after processing all answer sheets, the annotated answer sheets will be sent to student via email. As the algorithm is designed mainly to segment and decode handwritten student information and marks, the technical details of the finder pattern, student information recognition, optical mark recognition algorithms will be explained next.

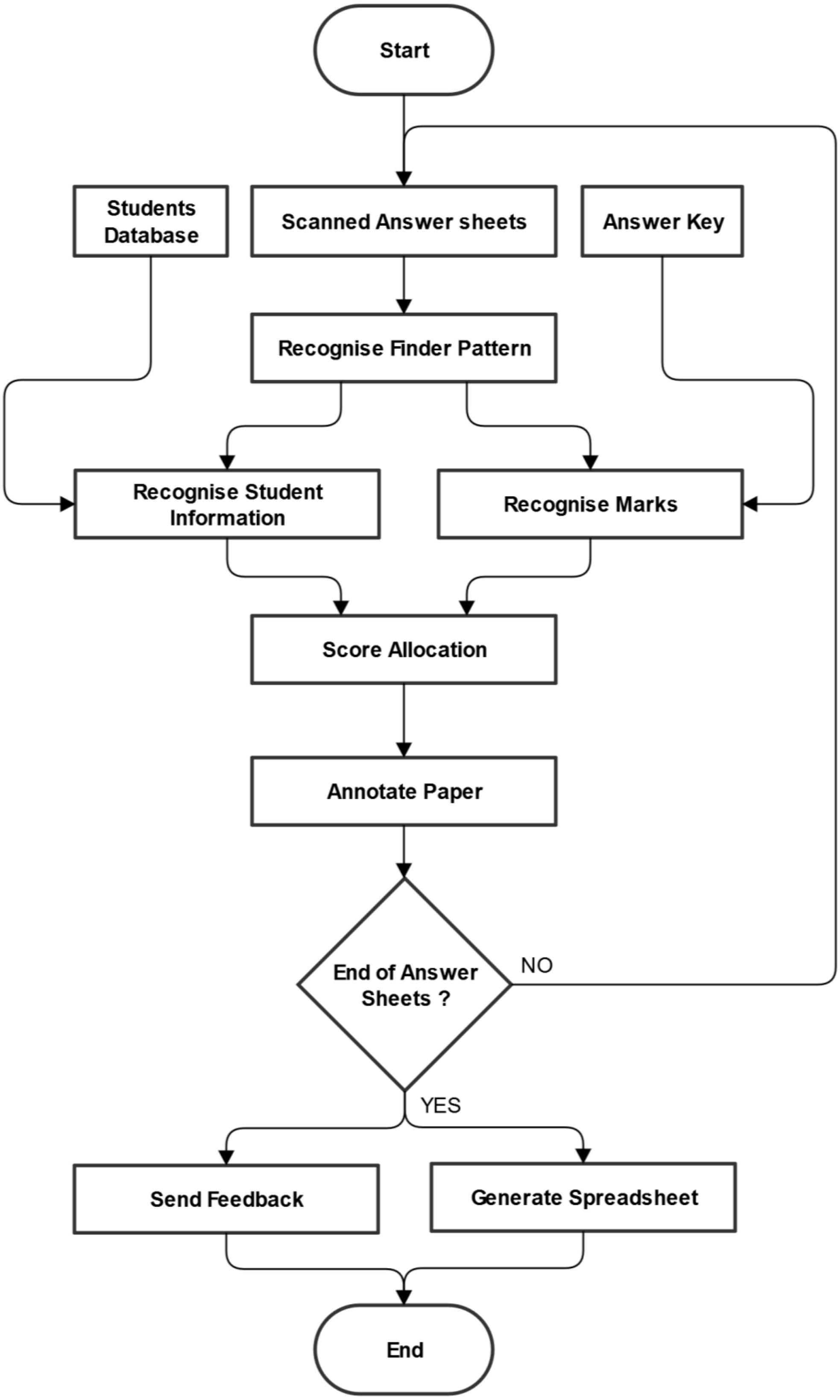


Figure 3.4: Flow chart of the proposed system.

3.3.1 Finder Pattern Recognition

The most essential part in the system flow is to register the answer sheet. By successfully registering the answer sheet, further processing can be applied such as student ID and mark recognition. However, a failure of registering the answer sheet means a failure in scoring the answer sheet. Unlike our eyes to recognise the shape of an answer sheet, computer software needs an algorithm to register an image. The aim of registering answer sheet is to define to the space of an answer sheet. It helps to know where to start and where to end which means every pixel in an answer sheet is registered, hence every pixel can be Located. There are several means to register an image in image processing. The complexity of image registration lies on the purpose and the conditions that are undertaken to register an image. Because the proposed system adopts an optical scanner to acquire the image, the complexity of registering the answer sheet is not high. Therefore, a reliable image registration method is adopted which finder pattern recognition.

Finder pattern technique has been used in various applications such as 1D barcodes and 2D barcodes. The complexity and the design of the pattern depend on the application. The proposed finder pattern is composed of three solid rectangles located in the margin of the answer sheet as shown in Figure 3.2 and Figure 3.3. The shape of the finder pattern is fixed and bold to be distinguishable by the algorithm. Every component's size is 21x8 pixel. The first component is located in the top left, the second component is in the top right, and the last component is located in the bottom left. As shown in Figure 3.5, the first component is the origin point as if in the Cartesian coordinate. The second component defines the maximum point in the vertical axis. The third point defines the maximum point in the horizontal axis of the answer sheet aspect ratio. The main roles of these three components are to define the landing point (origin point), to calculate the size of the scanned sheet, and to fix the answer sheet tilt.

Before locating the finder pattern, first the aspect ratio has to be checked. This aims to detect the answer sheet orientation if it was scanned vertically or horizontally. If it was scanned horizontally, the algorithm will rotate the image 90 degrees. Furthermore, if the answer sheet was scanned upside down, the algorithm looks for the logo which in the top left area of the answer sheet. If the algorithm does not see the pixel density of the logo, it rotates the answer sheet 180 degrees. Next, if the algorithm finds the logo, the algorithm is ready to locate the finder pattern.

To locate the finder pattern, first the image has to be converted into a binary image after applying a global threshold as shown in (2.16). The binary image is composed of either a black group of pixels or a white group of pixels, and because the size of the finder pattern points is pre-determined to 21x8 pixel, it is easy to erode every group of pixels that is greater or less than finder pattern points (21x8 pixel). After pixels erosion as shown in Figure 3.6, the image is left with three groups of pixels which are the finder pattern points.

Figure 3.5: The finder pattern design which highlights the three projection points location in the answer sheet, where P stands for point.

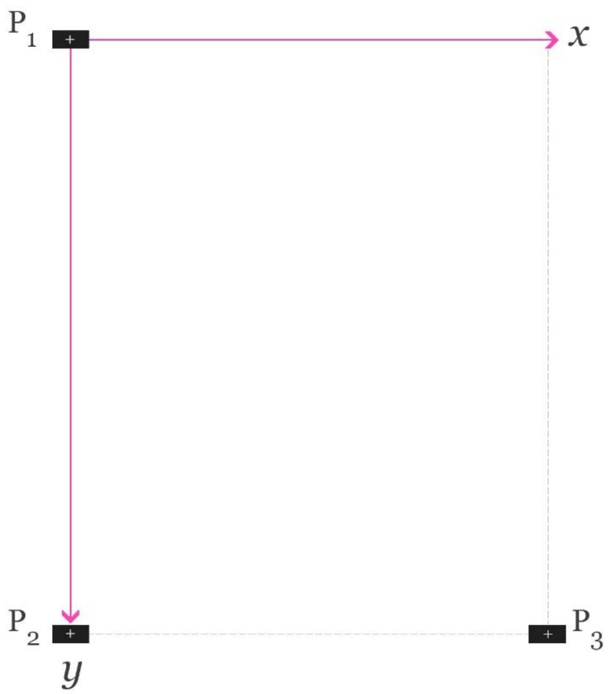


Figure 3.6: A concept of the answer sheet after pixel erosion.

Now, the centroids of the three pixel groups can be located. Points 1, 2, and 3 can be described as

$$P_1\,(x, y) = min(\,k(y)) \tag{3.1}$$

$$P_3(x, y) = max(\,k(x)) \;\cdot\; max(\,k(y)) \tag{3.2}$$

$$[P_2(x, y) > P_1\,(y)] \cap [P_2(x, y) < P_3\,(x)] \tag{3.3}$$

where $P_{n\,(\ldots)}$ is the centroid of the finder pattern point which has x and y value. $k(x, y)$ is 2×3 centroid matrix of the three points after the binarization which can be written as

$$k = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}. \tag{3.4}$$

As shown in (3.1), the first centroid with respect to y of the centroids set $k(x, y)$ is registered as the reference point of the image which is the first finder pattern point. Moreover, point 3 is located in the furthest location in x and y axis. Lastly, after locating points 1 and 3, point 2 must be greater than point 1 with respect to y or less than point 3 with respect to x.

After locating the exact location of the finder pattern, the size of the registered answer sheet can be estimated by

$$Size\ (x, y) = \left\{\sum_{i=P_1(x)}^{P_3(x)} x_i\ , \sum_{i=P_1(y)}^{P_2(y)} y_i\right\}. \quad (3.5)$$

As shown in (3.5), the width of the answer sheet is calculated by the sum of the pixels from the reference point until the centroid of the third point with respect to x. The height of the answer sheet is calculated by the sum of the pixels from the reference point until the centroid of the second point with respect to y.

After scanning the answer sheet, the answer sheet is not scanned precisely straight. The scanned image must exhibit a small angle of tilt clockwise or anti-clockwise as shown in Figure 3.7. From Figure 3.7, point 1 and 2 on the left side of the answer sheet are located in the same line. Therefore, as the location of both points are recognised, if there is a tilt, the angle θ of rotation between the two points can be estimated by taking the differences between the first point and the second which can be written as

$$\theta = \tan^{-1}\left[\frac{P_{1\,(x)} - P_{2\,(x)}}{P_{1\,(y)} - P_{2\,(y)}}\right], \quad (3.6)$$

where $P_{n\,(\ldots)}$ is the centroid of the finder pattern point with respect to x or y. Now, the image can be straightened by rotating the image by the same degree of rotation in the opposite direction. The corrected answer sheet $\grave{I}$ can be described as

$$\grave{I} = R\ I, \quad (3.7)$$

where I is the original answer sheet image, and R is the rotation matrix with the estimated rotation angle which can be written as

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \quad (3.8)$$

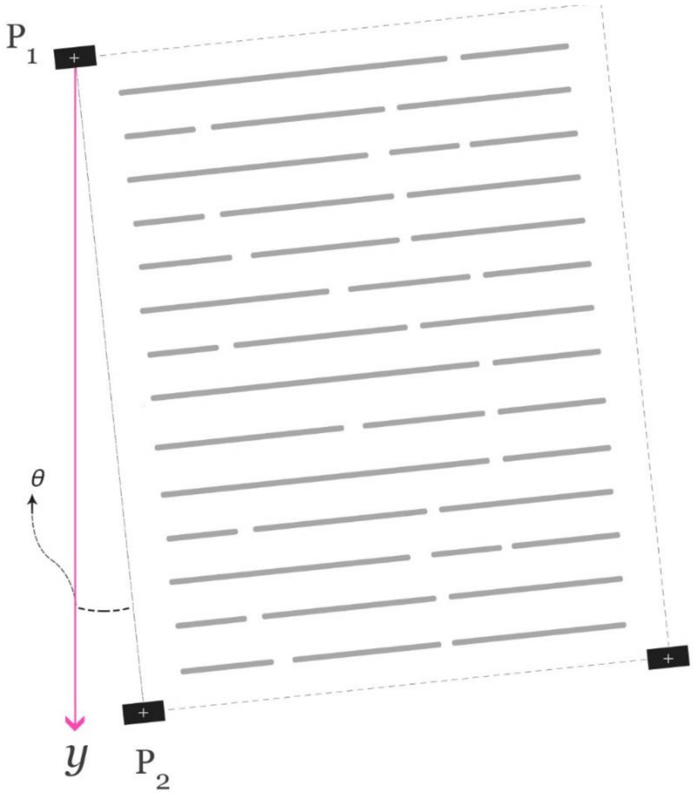


Figure 3.7: An example of answer sheet tilt during scanning, where θ is the angle of rotation.

After straightening the answer sheet, the student ID, name, and the answer area can be segmented. This is because the aspect ratios between the points and student ID and answer area are already determined. Now, every section of the answer sheet can be processed separately.

3.3.2 Student Information Recognition

Before detecting the overall score of an answer sheet, student's identity must be recognised. As shown in Figure 3.3, student area consists of two part; student ID and name. Most of the existing methods use the conventional student ID style which is called student ID matrix, to shade the numbers as shown in Figure 3.8. This method utilises only student ID and discards student name information. Moreover, it is found more restrictive and less intuitive for students to use if it is used for the first time. In this proposed approach, a novel method is designed to recognise student information in student area. The proposed student ID algorithm replaces the conventional style of encoding student IDs and utilises both student ID and name information for accurate and faster decoding results.

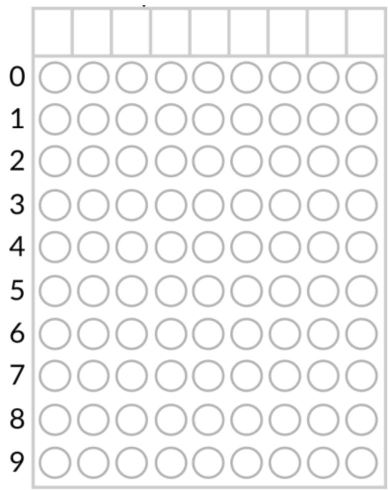


Figure 3.8: The conventional student ID recognition style [73].

The proposed algorithm is designed to decode student ID first and then recognise the number of the sur name characters. Later, the student ID with the right sur-name characters count is searched in the student database. This technique speeds up the search. However, if one or two characters of student ID are missed to decode or decoded wrongly, an auto-correction algorithm is designed to find the closest match with the existing database student IDs. Furthermore, a flag will be raised, so the user can intervene and validate the student information.

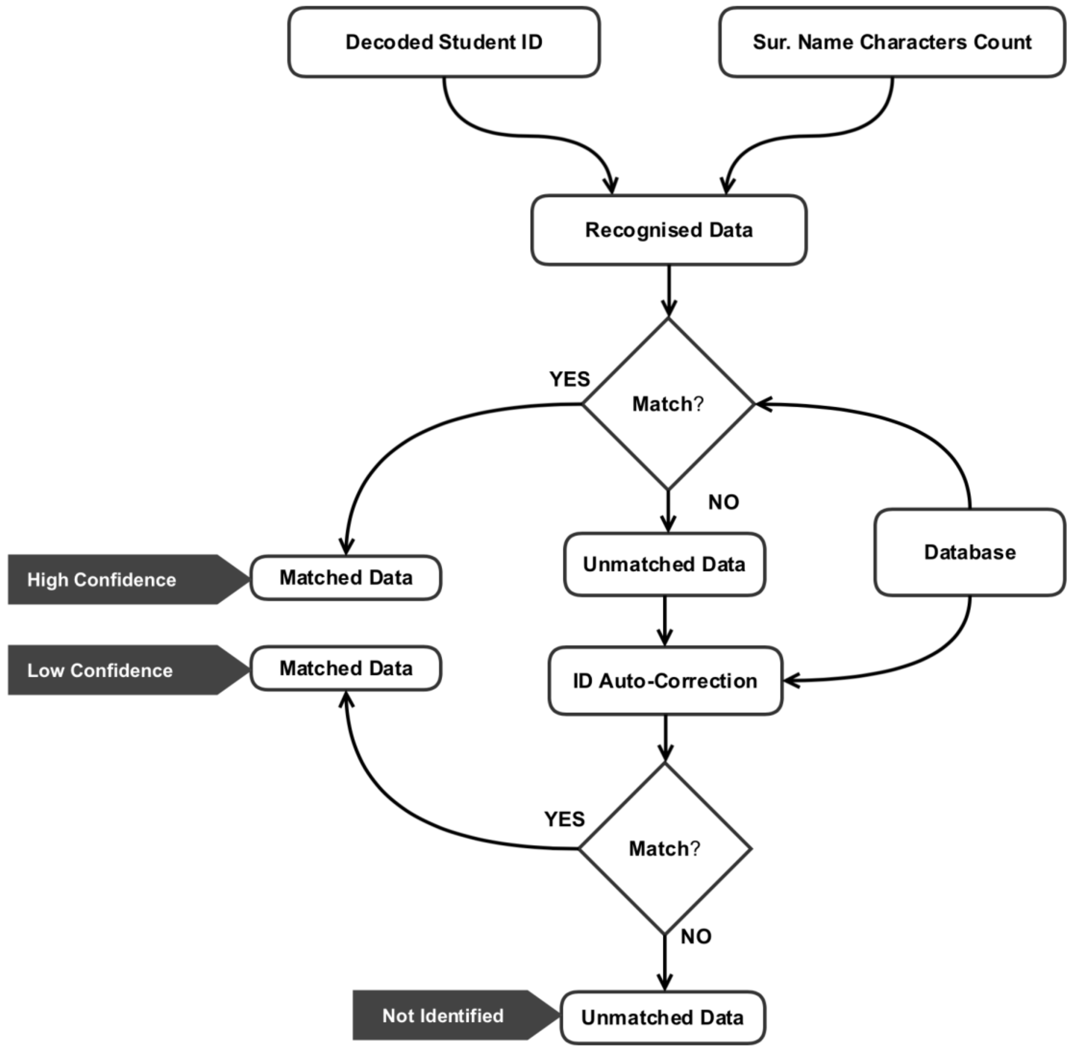


Figure 3.9: A flow chart of student information recognition algorithm.

Figure 3.9 shows the overall design of student information recognition algorithm. The inputs of the algorithm to start are a decoded ID from student ID area and the number of sur name characters. Both inputs are stored in a matrix as recognised data that need to be identified. Later, the algorithm looks for only the IDs that have the same sur name character count. If the ID is found, the ID will be stored as an Identified ID with a high confidence flag. However, if the ID is not found in the database, the unidentified ID will be passed to an auto correction algorithm as if one or more of the student ID character was decoded wrongly. Then, if the error was identified, corrected, and assigned to the right student name from the database, it will be flagged as an identified ID with low confidence as it has been corrected, and assigned to a wrong student. Finally, if the decoded ID could not be found in the database, it will be stored as it is with unidentified flag. Both data that are flagged with low confidence and not identified can be checked after the end of the whole scoring process.

Overall, this is the concept of the proposed student information recognition. Next, we are going to give more details about student ID recognition, name's character count recognition, and auto correction algorithms.

Student ID recognition algorithm is the soul of this section because student ID is the only unique identity for every test taker. Decoding the encoded ID successfully can guarantee student information to be found in database with high confidence. The recognition process is devised to make student ID encoding much more intuitive and less restrictive than the conventional method of most existing solutions (matrix identifier). However, a conventional handwritten character recognition was avoided because of the high complexity, and every student has a different handwriting style, hence it is almost impossible to recognise every student ID correctly. In this method, students have to write their IDs in segmented digits, just like a seven segment display. In this way, the possibility of facing irregular character shape is less, and the chances of writing style are limited and can be analysed accurately.

Figure 3.10 shows the design of student ID section which consists of up to eight segmented characters. In every character area, a template shows the segments pattern of how the character should be written on the dotted lines. The dotted lines are designed as a guideline for the test taker. In the instruction area, the steps on how to write an ID is explained briefly.

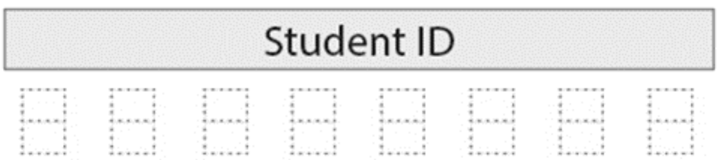


Figure 3.10: Student ID area in the answer sheet.

The algorithm is designed to interpret the numbers from zero to nine. Figure 3.11 shows the style of the handwriting of each number which indicates how easy to write a number. Students do not have to create strictly a straight line as shown in Figure 3.10.

The algorithm is designed to be more flexible with cures if they are within the region of interests (segments) as indicated in Figure 3.12 (in the seven blue areas).

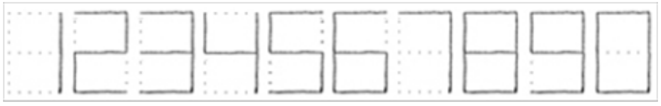


Figure 3.11: The style of the handwriting of all ten digits.

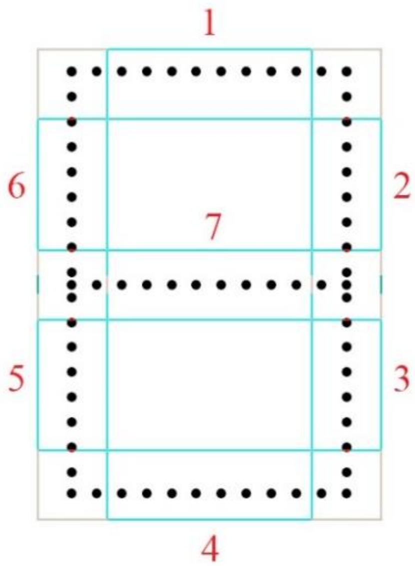


Figure 3.12: The seven segments active areas.

The algorithm starts by binarising and thinning (pre-processing) the characters to remove noise and the dotted lines by applying dilation and thinning operations. The aim is to retain only the connected pixel components. Then, there is a calculated threshold of pixel numbers for every segment. The threshold to be six pixels in each segment. If the number of pixels reaches the threshold, the segment is decoded as an active segment. After decoding every segment of a digit, the registered segments are multiplexed into a decimal number. Table 3.1 shows the multiplexing table of all the segments. Note that the segment numbering in Table 3.1 is the same numbering in Figure 3.12. The same steps are repeated for the rest of the digits. Finally, the decoded ID is stored as a recognised data for further processing as indicated in Figure 3.9 if needed. An example of number two recognition is illustrated in Figure 3.13. Further experiments on the robustness of the algorithm will be discussed in experimental results and discussion section.

Table 3.1: Active segments table.

Decimal	Active Segments Code						
	1	2	3	4	5	6	7
1	0	1	1	0	0	0	0
2	1	1	0	1	1	0	1
3	1	1	1	1	0	0	1
4	0	1	1	0	0	1	1
5	1	0	1	1	0	1	1
6	1	0	1	1	1	1	1
7	1	1	1	0	0	0	0
8	1	1	1	1	1	1	1
9	1	1	1	1	0	1	1
0	1	1	1	1	1	1	0

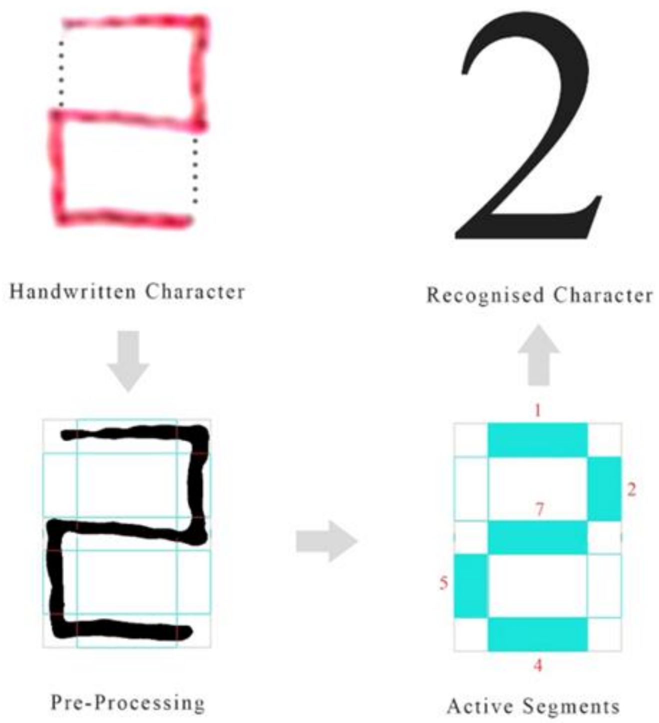


Figure 3.13: Student ID recognition steps.

After decoding student ID, the second input of student information recognition algorithm which is sur name character count will be discussed. As illustrated in Figure 3.14, this is designed student name area which contains letter boxes up to 16 boxes. The test taker has to write his name characters in the designated boxes. As alluded before, the aim of the letter separation is to attain the count of the name's letters. In this algorithm, the interest area is the sur name letters only. The technique is fairly simple to count letters in every box. Because the answer sheet is registered, every box can be reached and processed. Every letter box will be will be processed individually. Every box can hold approximately 360 pixels. Based on that, the number of box's pixels are checked. Threshold is set to be 100 pixels minimum. If the number of pixels is equal or greater than the threshold, a letter is counted. This happens to all the rest of the boxes. Then, the name characters' count is stored in recognised data as shown in Figure 3.9.

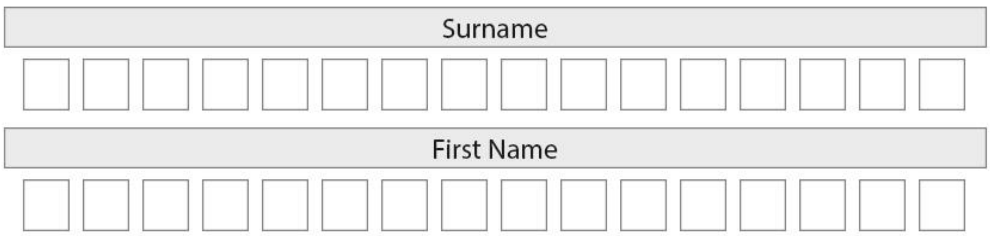


Figure 3.14: Student name section design.

After both algorithm's inputs have been decoded successfully, they will be matched with student' IDs and name. As shown in Figure 3.8, there are three possible outputs of the algorithm which are:

1- Matched ID with high confidence
2- Matched ID with low confidence
3- Not identified ID.

The first possible output occurs when a decoded ID is found in the database from the first matching time. The second possible output occurs when a decoded ID is not found in the database, but after applying auto correction algorithm. Lastly, the last possible outcome when the algorithm is not able to match the ID at after the first matching trial and applying auto correction. The auto correction algorithm is designed to minimise the matching error which may occur because faulty student ID recognition or faulty encoding by the test taker. The auto correction algorithm aims to predict and correct the wrong digit if only one or two digits of the eight digits were wrong. This is because more than two-digit correction may assign an ID to the wrong student.

The algorithm works simply by finding the wrong digits of every unmatched ID. This happens by applying the absolute error concept for every digit of the decoded ID with all the possible IDs in the existing database. The absolute error Δd can be estimated as

$$\Delta d = |d_0 - d|, \tag{3.9}$$

where d_0 is the unmatched ID digit, and d is every possible ID from database. The algorithm searches for the absolute error for each digit. The ID from database that has one or two absolutes errors is the best match to the unmatched ID. Next, the unmatched ID is replaced with the reference ID from database if the ID has not been assigned to another student. On the other hand, if the algorithm could not correct the unmatched ID, the ID will be stored as unidentified ID. By the end of the scoring process, the user can correct the ID manually. Figure 3.15 shows an example of a correction for a decoded ID with one wrong digit, where ID_d is the unmatched ID to be corrected, ID_{base} is the reference ID, Δ is the absolute error of each digit, and $ID_{matched}$ is the corrected ID.

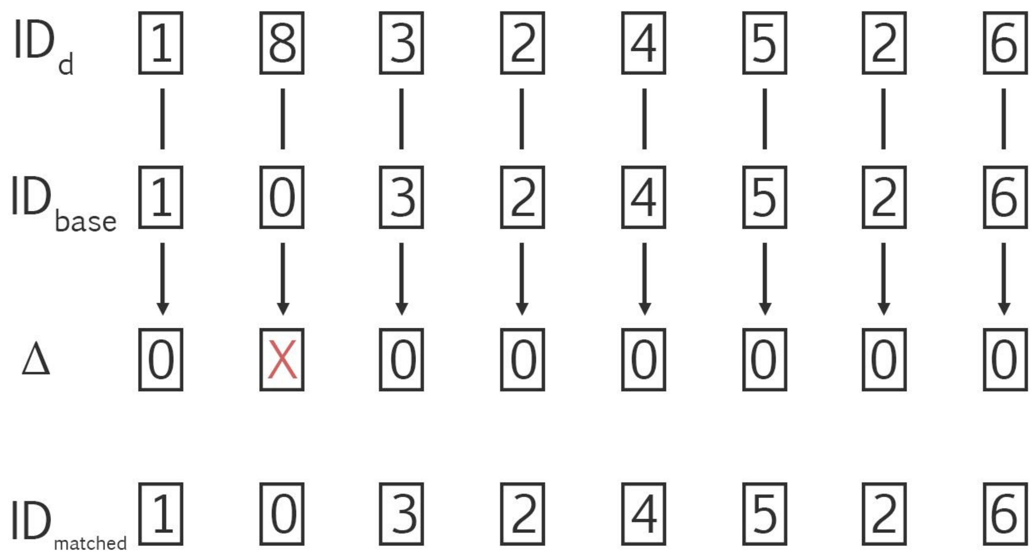


Figure 3.15: An example of an ID correction.

3.3.3 Marked Answer Recognition

Recognising mark marked answers is the core of this research. The proposed answer item is designed to handle five alternatives. In addition, the proposed optical mark recognition is designed to allow students to change their answers without a need to use a pencil and an eraser as shown in Figure 3.16, to change an answer, student needs to shade the area above the answer area and attempt again. Hence, the algorithm is designed to analyse the marks within these four possible cases for each answer item:

- No entry
- Single entry
- Multiple entries
- Single entry with changing answer attempt.

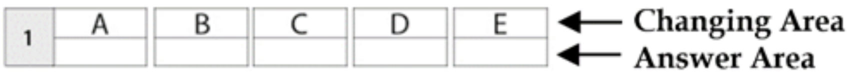


Figure 3.16: The proposed answer item, which contains five alternatives.

For the first case, where a student leaves the answer item with no attempt as shown in Figure 3.17(a). In the second case, a student answers only one alternative of the five alternatives as shown in Figure 3.17(b). In the third case, where a student answers multiple answers in the same answer item and no mark in the changing answer area as shown in Figure 3.17(c). For the fourth case, where a student answers one alternative and later changes the answer to another alternative. This leads to multiple entries in the answer area and one or more shadings in the changing answer area as shown in Figure 3.17(d). The proper method to shade the answer area is the same as what shown in Figure 3.17(b).

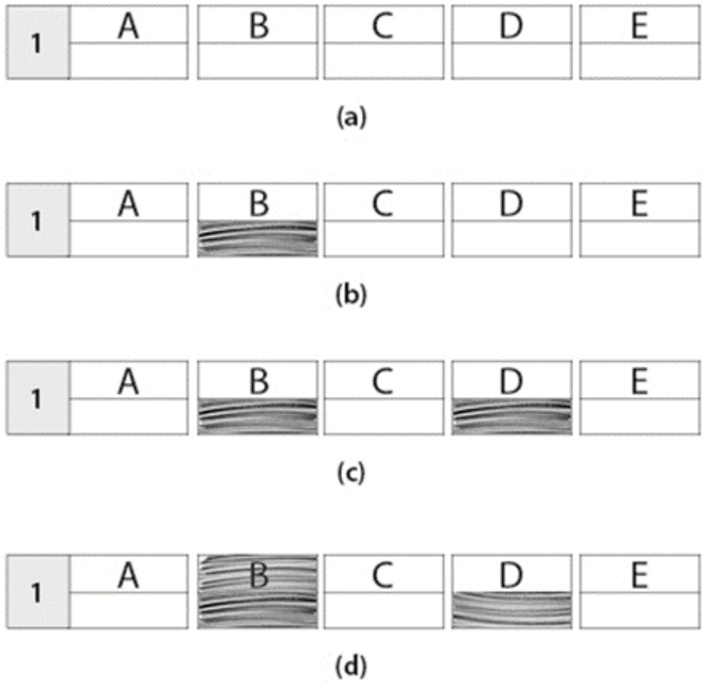


Figure 3.17: The four possible cases of answering the answer item, no entry (a) single entry (b), multiple entries (c), and single entry with changing attempt (d).

The proposed algorithm is designed based on the listed four possible cases. The block diagram of the proposed algorithm is shown in Figure 3.18, which handles multiple answer items. Before the algorithm runs, the key answer must be provided as reference to score each answer item. After recognising the three finder pattern points, the answer sheet is mapped. Hence, the exact location of each answer item is located. After that, starting from the left of answer item to check the first answer area (first answer box), the first answer box is checked by counting the number of black pixels and compare the number of pixels with a pre-calculated number of black pixels of a shaded box (threshold). The answer box is designed to have 2:1 aspect ratio. After resizing the answer sheet, the answer box size is 20x10 pixels, hence the total pixel number the answer box can contain is 200 pixels. Therefore, the threshold is set to be 80 pixel that is 40% of the answer box area $M(i,j)$ which can be defined as

$$M(i,j) = \begin{cases} 1, & B(x,y) > \mathrm{T} \\ 0, & B(x,y) < \mathrm{T} \end{cases}, \tag{3.10}$$

where $M(i,j)$ is a 5x72 matrix for the marked answers. The matrix has 5 columns as there are 5 answer boxes in every answer item. In addition, the matrix has 72 rows because there are 72 answer items in the answer sheet. $B(x, y)$ is the answer box area on the answer sheet. T is the threshold which equals to 80.

If the number of black pixels is less than the threshold, the marked answer is set to 0 in the matrix, and the algorithm will move to next answer box to check the threshold again until the last answer box. This means the first case has occurred, which is no entry case, and there will be no mark for this answer item. Moreover, if the number of black pixels are higher of the threshold, the area above the answer box area (changing answer area) will be checked. The method of checking the changing answer area is the same as (3.10) with a different threshold. Later, if yes, the Changing Answer Area is shaded based on the threshold. This indicates that the fourth possible case has occurred, which is single entry with changing answer attempt. The algorithm again will move to the next answer box area, and if not, the marked answer sheet will be checked with answer key as there is only a single entry (second case). If the marked answer matches the key answer, a score will be added to the overall score, and if not, the mark will be stay the same. However, the algorithm will check if there is another entry. If there is another entry of any of the answer box areas with a correct choice, this indicates that there are multiple entries (third case); hence, there will be no mark for this answer item. If all the answer boxes are checked, the algorithm will move to the second answer item, and iterates the same steps. The total score will be in a spreadsheet with the corresponding decoded student information.

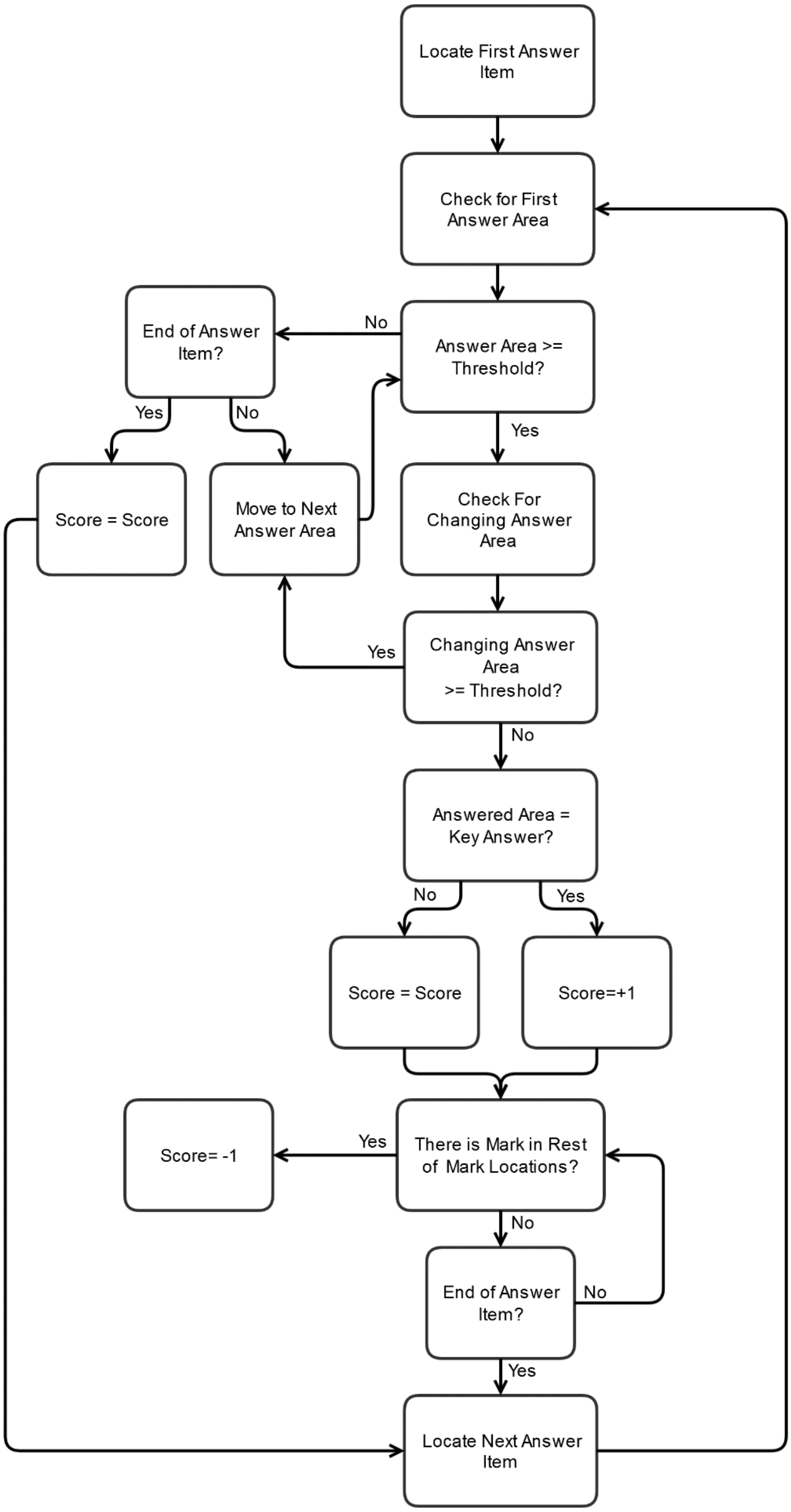


Figure 3.18: The block diagram of the proposed optical mark recognition algorithm.

3.3.4 Score Feedback

The final step of the system is to provide a feedback about the test. After scoring all answer items, the locations of the detected marks are stored to be annotated as shown in Figure 3.19. As illustrated in Figure 3.19, the shaded are-as are highlighted based the answer key. If the answer is correct, the area is highlighted green. In addition, if the answer is wrong, the area is highlighted red and corrected answer is highlighted blue. Finally, after scoring and annotating all answer sheets, the user has the option to check the results and send all sheets to students via email. The score feedback algorithm is a separate part of all the system. After the scoring process, the user can run the score feedback algorithm separately.

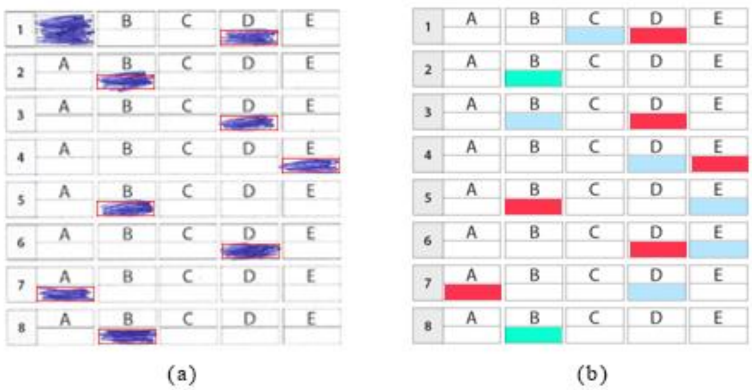


Figure 3.19: The annotation style, where red colour represents wrong answer, blue corrected answer and green correct answer.

3.4 Materials

The methodology requires hardware and software tools.

3.4.1 Hardware tools

As the research deals with optical data, hardware tools are required. First, the answer sheet needs to be printed and photocopied to many copies to be answered by students. Later, the user needs to scan the answers and feed them to the system; hence, the following hardware devices may be needed:

- Printer
- Photocopier
- Optical scanner
- Digital camera.

A printer is used to print the answer sheet. Then, a photocopier is utilised to photocopy the printed answer sheet to many copies as required for the assessment. After finishing the assessment, the answer sheets are transformed into optical data to electronic data via an optical scanner or a digital camera. For the purpose of speed and clarity, an optical scanner is recommended over a digital camera as it is designed to acquire documents. That's because scanners are designed specially more for acquiring documents Cameras are not designed capture documents. They designed to capture photos in general. Nonetheless, a digital camera can be used if answer sheets need to add to the system individually.

Moreover, there is another a professional hardware fits the need of the proposed system, which is a multifunction device (all in one device). A multifunction device is a professional office device that provides a range of functionalities for office use. The multifunction device contains a printer, photocopier, optical scanner, and a fax in one device. There are many models of it in the market. It comes in different sizes, printing and scanning speed, and printing and scanning capacity [74]. The multifunction device that will be used for the proposed method is Aficio MP C5501A in ECU Engineering School. Aficio MP C5501A has many features suit the need of the proposed method. The main feature is that Aficio MP C5501A can copy 55 answer sheets in a minute. Moreover, it can scan multiple answer sheets once from the document feeder, not like the ordinary scanner where the user has to scan all answer sheets individually [75].

3.4.2 Software tool

The proposed method can be implemented via several software tools and computer languages such as OpenCV, Matlab Image Processing tool, and C++ programming language. However, some of image processing tools and computer languages require high level of programming proficiency. Furthermore, some of the tools and programming languages are not image processing solving problem oriented. One of the best option to implement is Matlab Image Processing Tool for several reasons. Matlab is a software tool fully designed for engineers and scientists who do not have to the professional programming skills to develop their projects. Matlab brings the most and easiest environment to develop algorithms and models. Moreover, it combines a desktop environment with high level programming languages such as C, Java, and Python [76]. More importantly, Matlab software provides an image processing tool to develop document analysis projects. Matlab Image Processing ToolBox offers a variety of set of references for algorithms, functions, and applications for image processing. The tool will be useful and suitable for the proposed method as it can perform image enchantment, segmentation, and registration [67]. Besides MATLAB, Adobe Photoshop CC was utilised to design the answer sheet entirely.

4. RESULTS AND DISCUSSION

In the fourth chapter, various aspects of the proposed system will be discussed. This chapter starts by discussing the stages of the answer sheet design in Section 4.1. Later, Section 4.2 will present the experimental setup. An analysis of software algorithm will be carried out in Section 4.3. Then, the accuracy of the software algorithm will be discussed in Section 4.4. In Section 4.5, the speed of the system will be discussed. Results of a case study will be discussed in Section 4.6. In Section 4.7, the deduced limitation of the system will be listed.

4.1 Outcome of Answer Sheet Design

The answer sheet design has been enhanced throughout the system design development stages. The principle of answer sheet designs is set to meet the answer sheet design layout as shown in Figure 3.2. There were two major answer sheet versions were proposed. Both designs are slightly similar to serve the purpose of the use. However, both have differences in the details.

The answer sheet version which is shown in Figure 4.1 was the first answer sheet design. As shown in the figure, the student information utilises student ID matrix as the main input to recognise student information. Moreover, in answer area, the shading boxes is small which adopts underlining instead of shading. Because of the small answer box area design, the answer sheet can handle up to 99 answer items. Furthermore, the answer sheet was design to have only two finder pattern points.

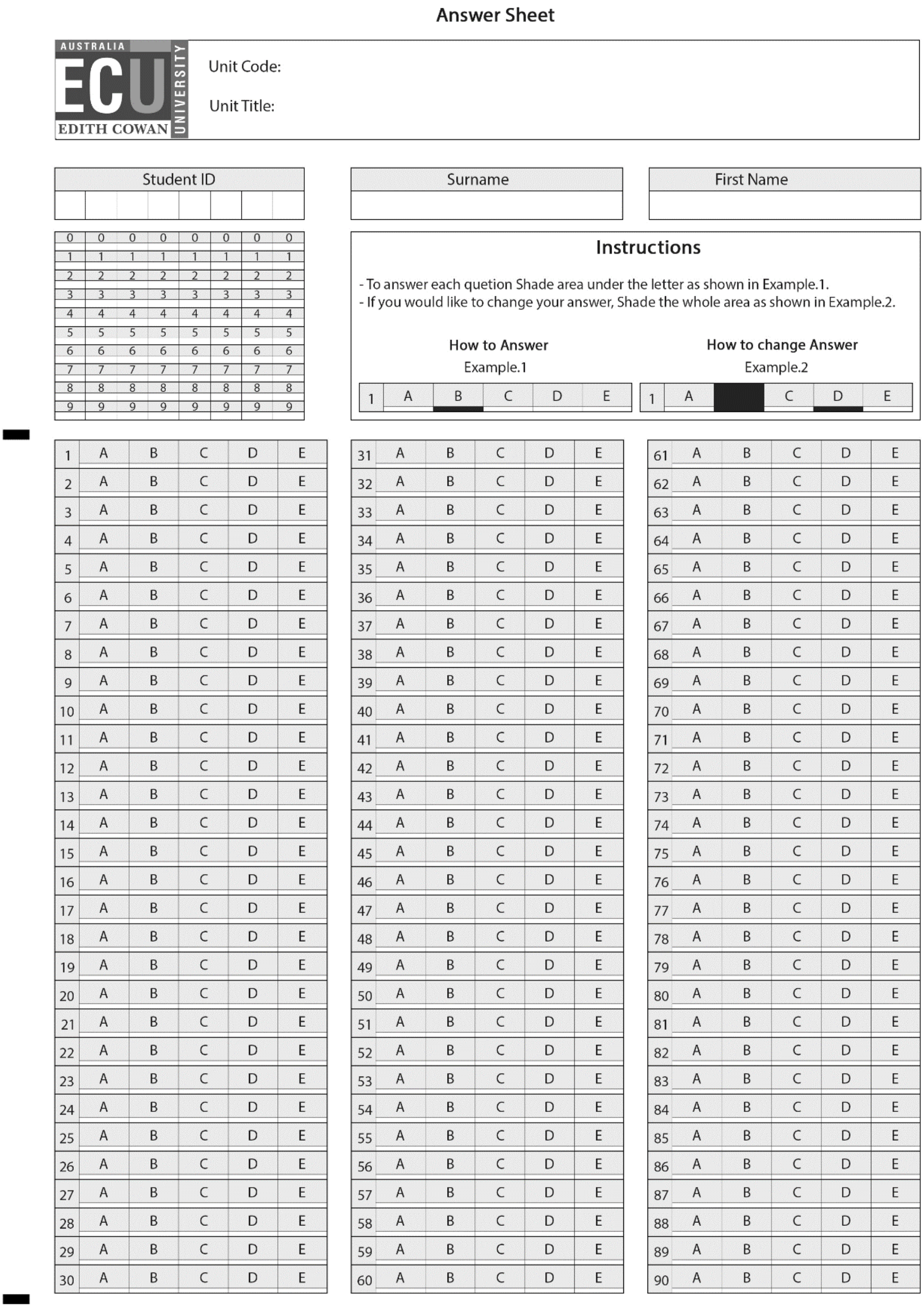
However, after many individual testing, it was figured out that:

- As the answer box is very thin, the thing design made it more difficult to recognise the marked area which affects the marked answer algorithm accuracy negatively. Based on the setup local threshold, some of the strokes lines in the answer area were recognised as marked answers. In addition, if the answers were marked by a light pen colour such as a pensile, some of the marked answers disappear after binarisation and noise reduction process. Therefore, the answer boxes needed to be optimised to improve the accuracy of the marked answer detection.

- The finder pattern points were designed to be the reference points of the sheet. However, the position the first finder pattern point needed to be reconsidered as it lies under the instruction area which makes it easier to segment answer area, but it makes difficult to segment student information area because the height above the first finder pattern point cannot be determined. Despite the position of the finder patterns, the number of the finder pattern points needs to be reconsidered as well. Two finder pattern points in the left side of the answer sheet are just enough to correct the answer sheet tilt after scanning. In addition, they help to determine the height of the answer sheet. Nevertheless, it was investigated that a third point is required to determine to the width as well after reposition the first finder pattern point.

Hence, the answer sheet had to have an overhaul to fix the mentioned issues. The answer sheet version 2.0 is shown in Figure 4.2. As shown in the figure:

- ✓ The answer boxes were enlarged to 300% of the last answer box area which helps to improve the accuracy of the marked answer detection after applying the local threshold. Moreover, the answer box area optimisation helped to avoid any confusion with adjacent strokes lines after binarisation and noise reduction process. However, because the answer box areas have been enlarged, the number of the answer items has decrease to 72 instead of 90 items.
- ✓ The first finder pattern point was repositioned to cover the student information area. In addition, a third finder pattern was added to resolve the sheet width issue. Moreover, the size of the finder pattern points was adjusted to be slightly bigger which eases the finder pattern points detection as the answer box area was enlarged.
- ✓ Lastly, all the answer sheet sections were refined to have a lighter design. This gives the answer sheet better readability, and the reduces the amount of the printing ink as the answer sheet may be printed with a large amount of copies.



Answer Sheet

AUSTRALIA ECU EDITH COWAN UNIVERSITY

Unit Code:

Unit Title:

Student ID	Surname	First Name

0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9

Instructions

- To answer each quetion Shade area under the letter as shown in Example.1.
- If you would like to change your answer, Shade the whole area as shown in Example.2.

How to Answer

Example.1

1	A	B	C	D	E

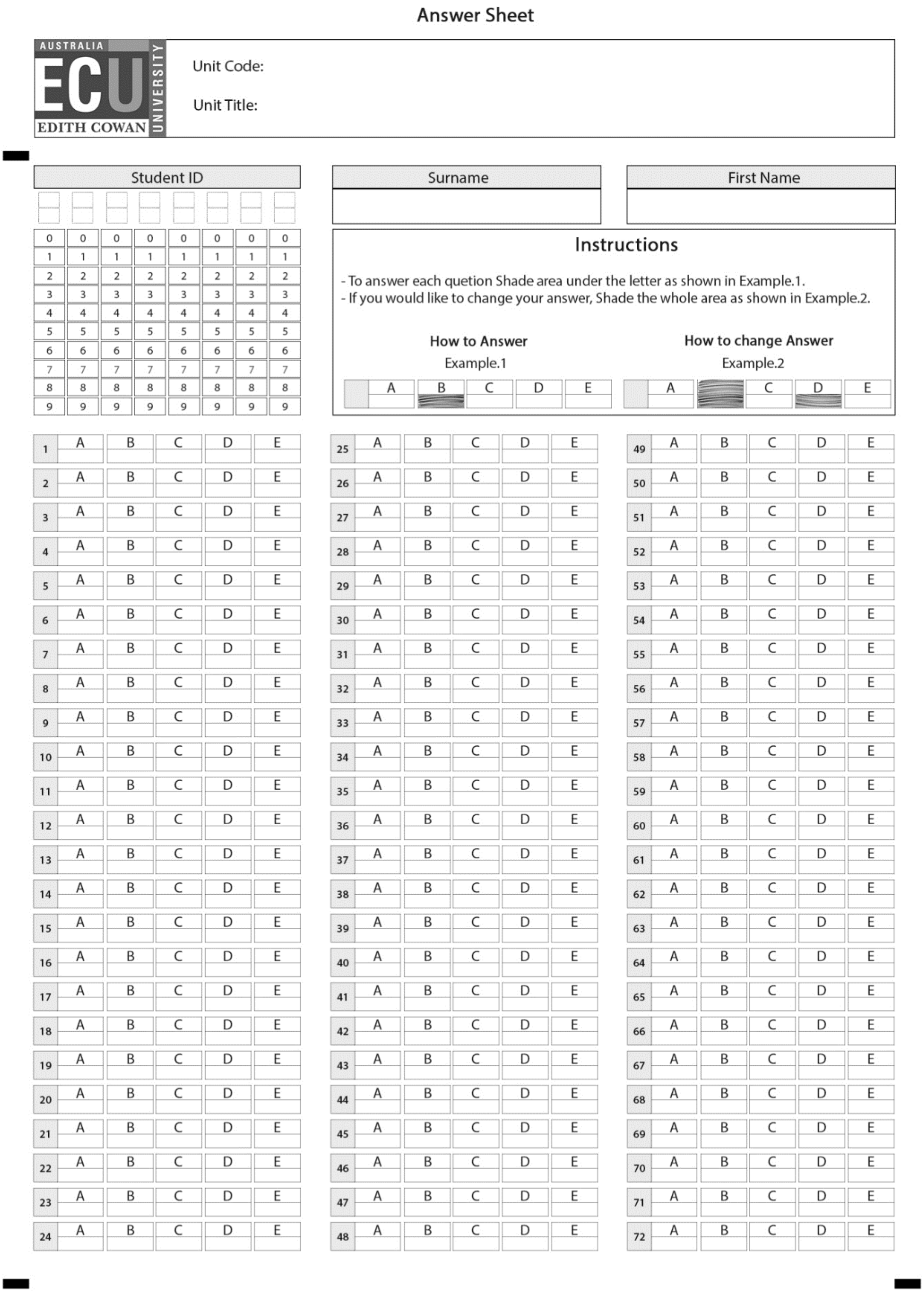
How to change Answer

Example.2

1	A		C	D	E

1	A	B	C	D	E	31	A	B	C	D	E	61	A	B	C	D	E
2	A	B	C	D	E	32	A	B	C	D	E	62	A	B	C	D	E
3	A	B	C	D	E	33	A	B	C	D	E	63	A	B	C	D	E
4	A	B	C	D	E	34	A	B	C	D	E	64	A	B	C	D	E
5	A	B	C	D	E	35	A	B	C	D	E	65	A	B	C	D	E
6	A	B	C	D	E	36	A	B	C	D	E	66	A	B	C	D	E
7	A	B	C	D	E	37	A	B	C	D	E	67	A	B	C	D	E
8	A	B	C	D	E	38	A	B	C	D	E	68	A	B	C	D	E
9	A	B	C	D	E	39	A	B	C	D	E	69	A	B	C	D	E
10	A	B	C	D	E	40	A	B	C	D	E	70	A	B	C	D	E
11	A	B	C	D	E	41	A	B	C	D	E	71	A	B	C	D	E
12	A	B	C	D	E	42	A	B	C	D	E	72	A	B	C	D	E
13	A	B	C	D	E	43	A	B	C	D	E	73	A	B	C	D	E
14	A	B	C	D	E	44	A	B	C	D	E	74	A	B	C	D	E
15	A	B	C	D	E	45	A	B	C	D	E	75	A	B	C	D	E
16	A	B	C	D	E	46	A	B	C	D	E	76	A	B	C	D	E
17	A	B	C	D	E	47	A	B	C	D	E	77	A	B	C	D	E
18	A	B	C	D	E	48	A	B	C	D	E	78	A	B	C	D	E
19	A	B	C	D	E	49	A	B	C	D	E	79	A	B	C	D	E
20	A	B	C	D	E	50	A	B	C	D	E	80	A	B	C	D	E
21	A	B	C	D	E	51	A	B	C	D	E	81	A	B	C	D	E
22	A	B	C	D	E	52	A	B	C	D	E	82	A	B	C	D	E
23	A	B	C	D	E	53	A	B	C	D	E	83	A	B	C	D	E
24	A	B	C	D	E	54	A	B	C	D	E	84	A	B	C	D	E
25	A	B	C	D	E	55	A	B	C	D	E	85	A	B	C	D	E
26	A	B	C	D	E	56	A	B	C	D	E	86	A	B	C	D	E
27	A	B	C	D	E	57	A	B	C	D	E	87	A	B	C	D	E
28	A	B	C	D	E	58	A	B	C	D	E	88	A	B	C	D	E
29	A	B	C	D	E	59	A	B	C	D	E	89	A	B	C	D	E
30	A	B	C	D	E	60	A	B	C	D	E	90	A	B	C	D	E

Figure 4.1: Answer sheet version 1.0.



Answer Sheet

AUSTRALIA ECU EDITH COWAN UNIVERSITY

Unit Code:

Unit Title:

Student ID							
0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9

Surname

First Name

Instructions

- To answer each quetion Shade area under the letter as shown in Example.1.
- If you would like to change your answer, Shade the whole area as shown in Example.2.

How to Answer
Example.1

A B C D E

How to change Answer
Example.2

A B C D E

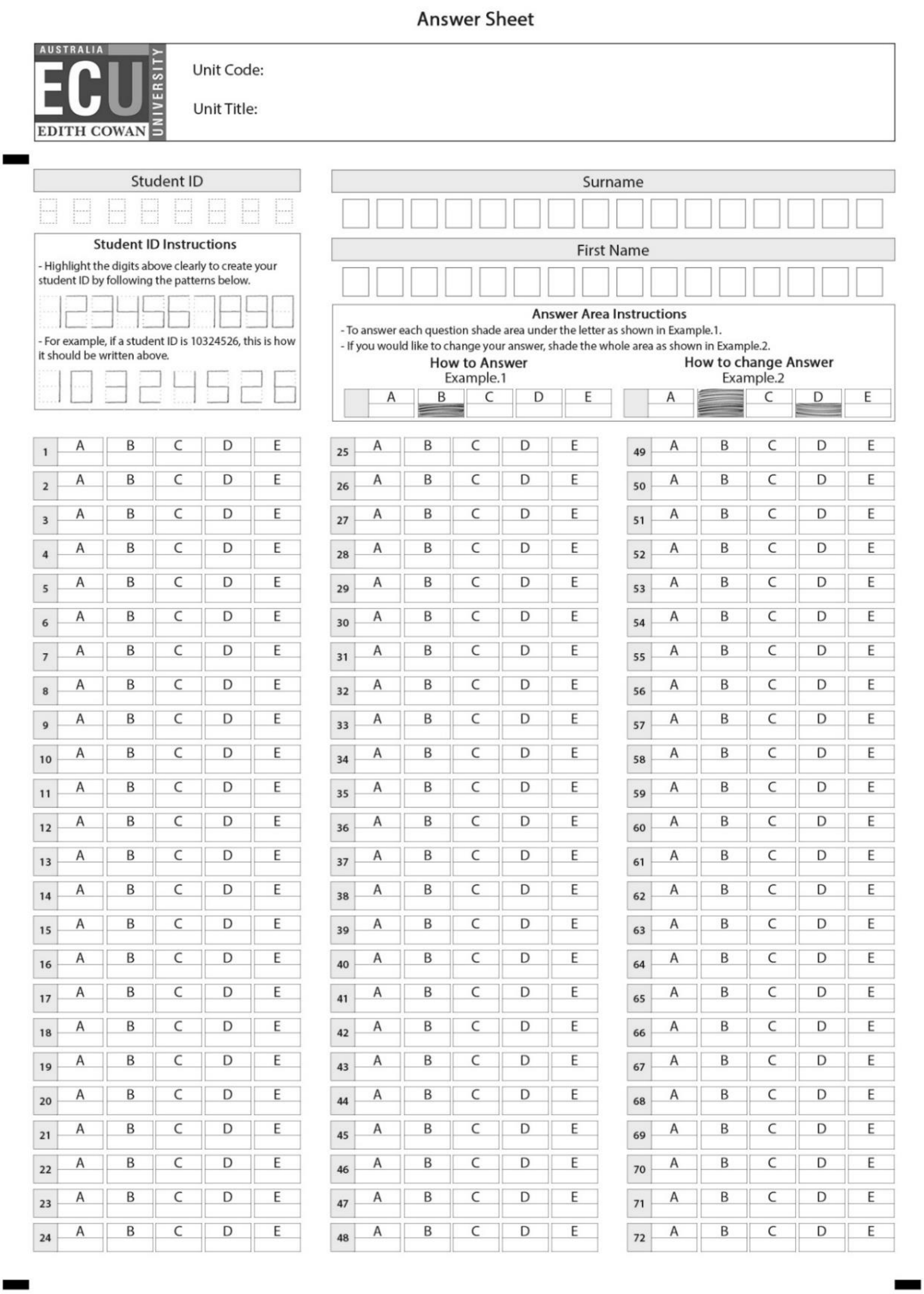
1	A	B	C	D	E	25	A	B	C	D	E	49	A	B	C	D	E
2	A	B	C	D	E	26	A	B	C	D	E	50	A	B	C	D	E
3	A	B	C	D	E	27	A	B	C	D	E	51	A	B	C	D	E
4	A	B	C	D	E	28	A	B	C	D	E	52	A	B	C	D	E
5	A	B	C	D	E	29	A	B	C	D	E	53	A	B	C	D	E
6	A	B	C	D	E	30	A	B	C	D	E	54	A	B	C	D	E
7	A	B	C	D	E	31	A	B	C	D	E	55	A	B	C	D	E
8	A	B	C	D	E	32	A	B	C	D	E	56	A	B	C	D	E
9	A	B	C	D	E	33	A	B	C	D	E	57	A	B	C	D	E
10	A	B	C	D	E	34	A	B	C	D	E	58	A	B	C	D	E
11	A	B	C	D	E	35	A	B	C	D	E	59	A	B	C	D	E
12	A	B	C	D	E	36	A	B	C	D	E	60	A	B	C	D	E
13	A	B	C	D	E	37	A	B	C	D	E	61	A	B	C	D	E
14	A	B	C	D	E	38	A	B	C	D	E	62	A	B	C	D	E
15	A	B	C	D	E	39	A	B	C	D	E	63	A	B	C	D	E
16	A	B	C	D	E	40	A	B	C	D	E	64	A	B	C	D	E
17	A	B	C	D	E	41	A	B	C	D	E	65	A	B	C	D	E
18	A	B	C	D	E	42	A	B	C	D	E	66	A	B	C	D	E
19	A	B	C	D	E	43	A	B	C	D	E	67	A	B	C	D	E
20	A	B	C	D	E	44	A	B	C	D	E	68	A	B	C	D	E
21	A	B	C	D	E	45	A	B	C	D	E	69	A	B	C	D	E
22	A	B	C	D	E	46	A	B	C	D	E	70	A	B	C	D	E
23	A	B	C	D	E	47	A	B	C	D	E	71	A	B	C	D	E
24	A	B	C	D	E	48	A	B	C	D	E	72	A	B	C	D	E

Figure 4.2: Answer sheet version 2.0.

As the research aims to find a novel method to decode student information, the second answer sheet was modified in the student information section to meet the novel student information recognition algorithm. The modified answer sheet is shown in Figure 4.3. The modifications are as follow:

- ✓ The first name and surname boxes were redesigned to segment each name character. This helps to count the number of characters in student name area as mentioned in Section 3.2.2.
- ✓ The student ID boxes were redesigned to suite the requirements the novel student information recognition algorithm. In addition, the student ID matrix was dropped and replace student ID instruction.

The modified answer sheet is the final answer sheet design. Based on this answer sheet, the individual experiments and the case study were tested. And the results of all software algorithms will be discussed in the following sub-sections.



Answer Sheet

AUSTRALIA ECU EDITH COWAN UNIVERSITY

Unit Code:

Unit Title:

Student ID

Student ID Instructions

- Highlight the digits above clearly to create your student ID by following the patterns below.

1234567890

- For example, if a student ID is 10324526, this is how it should be written above.

10324526

Surname

First Name

Answer Area Instructions

- To answer each question shade area under the letter as shown in Example.1.
- If you would like to change your answer, shade the whole area as shown in Example.2.

How to Answer
Example.1

A B C D E

How to change Answer
Example.2

A B C D E

1	A	B	C	D	E	25	A	B	C	D	E	49	A	B	C	D	E
2	A	B	C	D	E	26	A	B	C	D	E	50	A	B	C	D	E
3	A	B	C	D	E	27	A	B	C	D	E	51	A	B	C	D	E
4	A	B	C	D	E	28	A	B	C	D	E	52	A	B	C	D	E
5	A	B	C	D	E	29	A	B	C	D	E	53	A	B	C	D	E
6	A	B	C	D	E	30	A	B	C	D	E	54	A	B	C	D	E
7	A	B	C	D	E	31	A	B	C	D	E	55	A	B	C	D	E
8	A	B	C	D	E	32	A	B	C	D	E	56	A	B	C	D	E
9	A	B	C	D	E	33	A	B	C	D	E	57	A	B	C	D	E
10	A	B	C	D	E	34	A	B	C	D	E	58	A	B	C	D	E
11	A	B	C	D	E	35	A	B	C	D	E	59	A	B	C	D	E
12	A	B	C	D	E	36	A	B	C	D	E	60	A	B	C	D	E
13	A	B	C	D	E	37	A	B	C	D	E	61	A	B	C	D	E
14	A	B	C	D	E	38	A	B	C	D	E	62	A	B	C	D	E
15	A	B	C	D	E	39	A	B	C	D	E	63	A	B	C	D	E
16	A	B	C	D	E	40	A	B	C	D	E	64	A	B	C	D	E
17	A	B	C	D	E	41	A	B	C	D	E	65	A	B	C	D	E
18	A	B	C	D	E	42	A	B	C	D	E	66	A	B	C	D	E
19	A	B	C	D	E	43	A	B	C	D	E	67	A	B	C	D	E
20	A	B	C	D	E	44	A	B	C	D	E	68	A	B	C	D	E
21	A	B	C	D	E	45	A	B	C	D	E	69	A	B	C	D	E
22	A	B	C	D	E	46	A	B	C	D	E	70	A	B	C	D	E
23	A	B	C	D	E	47	A	B	C	D	E	71	A	B	C	D	E
24	A	B	C	D	E	48	A	B	C	D	E	72	A	B	C	D	E

Figure 4.3: Answer sheet version 2.1.

4.2 Experimental Setup

To experiment the proposed system, an experimental setup is needed. The experimental setup consists of hardware and software tools. These tools were utilised to experiment the proposed system. The tools functions involve:

- Printing
- Scanning
- processing and storing data.

Printing and scanning functions are related only to the answer sheet. Whereas, the processing and storing functions are related to the software algorithm which are crucial function to run and test the algorithm. For both printing and scanning function, the multifunction device Ricoh Aficio MP C5501A was utilised which can print and scan multiple answer sheets as discussed in Section 3.4. Furthermore, the specifications of the processing and storing tool are:

- Computer model: MacBook Air
- Computer hardware specs: Intel Core i7 2.4 GHz processor with 8GB RAM
- Operating system: MAC OS X Yosemite 10.10
- Programming language: MATLAB R2012a
- Database format: Microsoft Excel 2016.

All the testing results in this chapter were obtained by these tools.

4.3 Software Algorithm Analysis

In this section, the software algorithm will be analysed. A step by step analysis of the algorithm process will be shown of an individual answer sheet. The answer sheet that will be used in this analysis is shown in Figure 4.4. All the experimental setup tools were used for the example.

After printing and filling up the answer sheet, the answer sheet was scanned via Aficio MP C5501A. The scanned image is as shown in Figure 4.4. Then, in MATLAB, the three inputs of the systems which are answer key excel file, student database excel file, and the directory of the answer sheet image have to be provided to run the algorithm script.

After providing the inputs, the algorithm first resizes the answer sheet to 1000 pixel of the longest aspect ratio with 75 pixels per inch. Hence, the image size is 1000X713 pixel. Reducing the size of the answer sheet is important to speed up the processing time. And the reason behind 1000-pixel reduction is that 1000-pixel answer sheet is readable and small enough for fast processing.

After resizing, the answer sheet needs to be clean for finder pattern detection. Finder pattern detection is important to correct image tilt during scanning and map the whole answer sheet. And to do that, the answer sheet need to be enhanced to erode the unwanted noise. Hence, every pixel group in the image (noise) is unwanted at this stage except the finder pattern points. To erode the noise, first, the image is converted to a grayscale. Then a local binarisation threshold is used to end up with a binary image. The result of this step is shown in Figure 4.5. Next, pixel groups need to be filtered to detect the finder pattern points. Pixel group separation can be achieved as the size of the finder pattern points are known which is 21x8 pixels. To have only the finder pattern points only in the binary image, every pixel group size in Figure 4.6 (a) smaller than the finder pattern points is filtered out as shown in Figure 4.6 (a). As it can be seen in the figure, the image contains only the finder pattern points and the logo which is bigger than the finder pattern. Next, the logo needs to be separated. The logo can be separated by filtering out any pixel smaller or equal to the finder pattern points as shown in Figure 4.6 (b). Now by taking the difference between Figure 4.6 (a) and 4.6 (b), the resultant image is an image with the three finder pattern points only as shown in Figure 4.7.

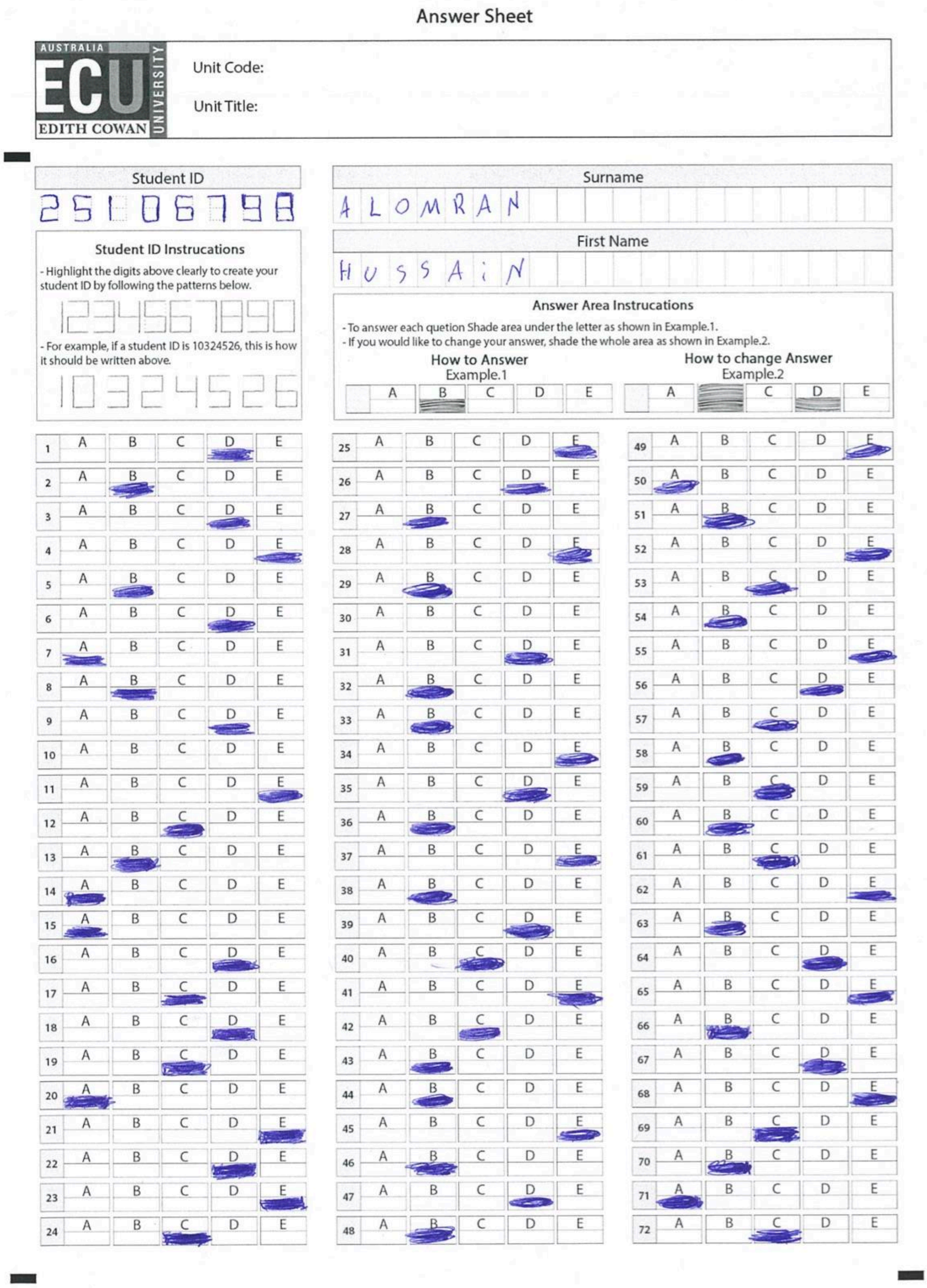


Figure 4.4: The scanned answer sheet.

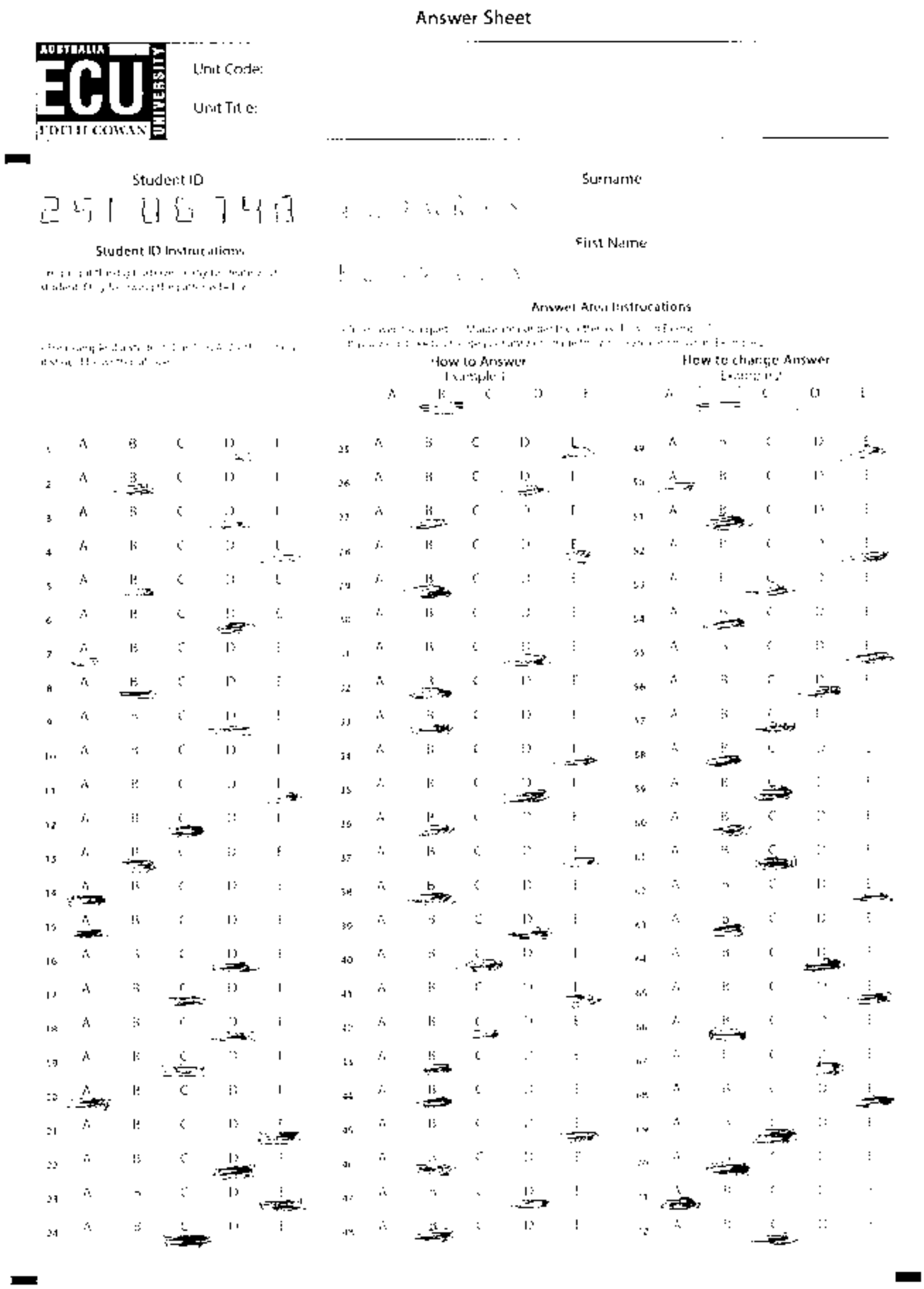
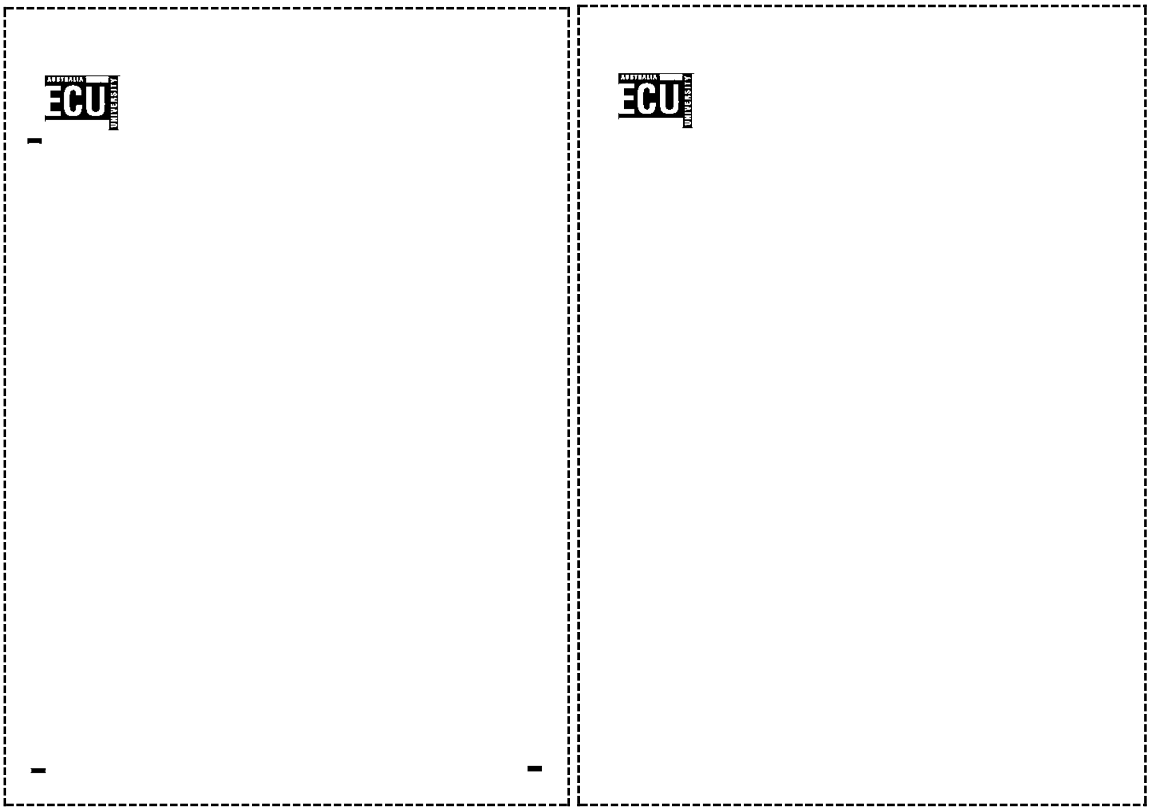


Figure 4.5: The binarised answer sheet.



(a) (b)

Figure 4.6: Pixel group separation, where (a) showed the filtered group pixel smaller than the finder pattern, and (b) shows the separated pixel group bigger than the finder pattern.

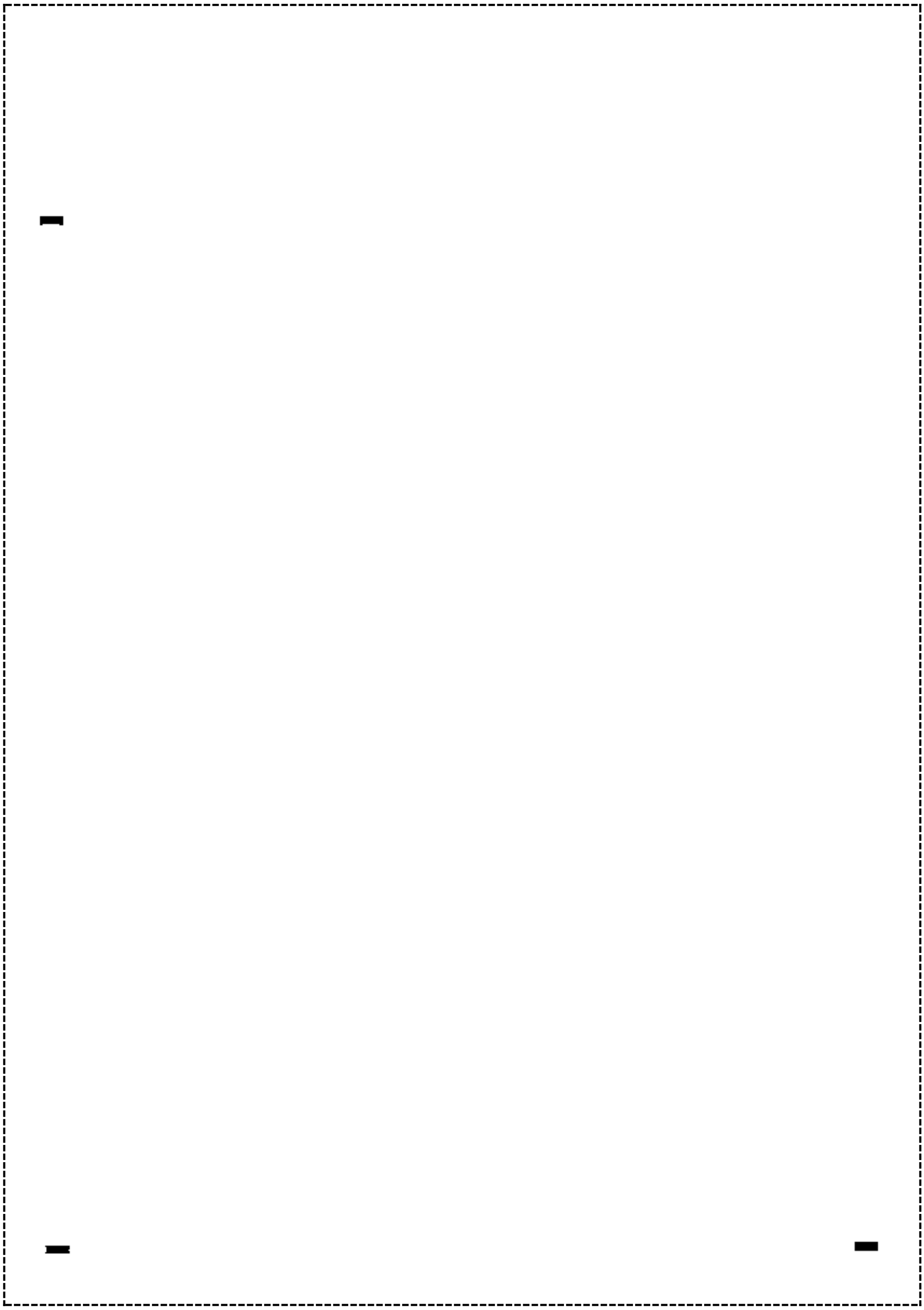


Figure 4.7: The filtered image which is the result of the difference between Figure 4.6 (a) and Figure 4.6 (b).

Now after filtering the noise of the image, the remaining pixel groups are the three finder pattern points. This makes it easier to locate the three points. The closest pixel group to the origin of the image coordinate with respect to the X axis is the first finder pattern point. And the furthest pixel group with respect to Y axis is the third point. Now, the remaining pixel group must be the second finder point. The registered centroids of the finder pattern points are shown in Table 4.1. By registering the three points, the point locations can be referred to the original image (Figure 4.4).

Table 4.1: The centroids of finder pattern on the image.

Finder Pattern	X value	Y value
Point 1	36.557	165.655
Point 2	41.000	958.500
Point 3	666.568	959.976

The point locations can be utilised to estimate the image tilt and correct it. By using (3.6). By evaluate the values in the equation, the result is:

$$\theta = \tan^{-1}\left[\frac{36.557-41}{165.655-958.5}\right] = -0.32^{\circ}. \quad (4.1)$$

As it is evaluated for the equation, the answer sheet is tilted by 0.32 degree anticlockwise. Therefore, the tilt can be corrected by compensating 0.32 degree clockwise which will make the answer sheet straight. After tilt correction, the image is segmented to include only the interest regions only which are student information and answer area. The cropped image is shown in Figure 4.8. From the figure, the image is cropped from the first finder pattern point to the third point. The positions of the interest regions are pre-calculated and can be segmented to be process individually.

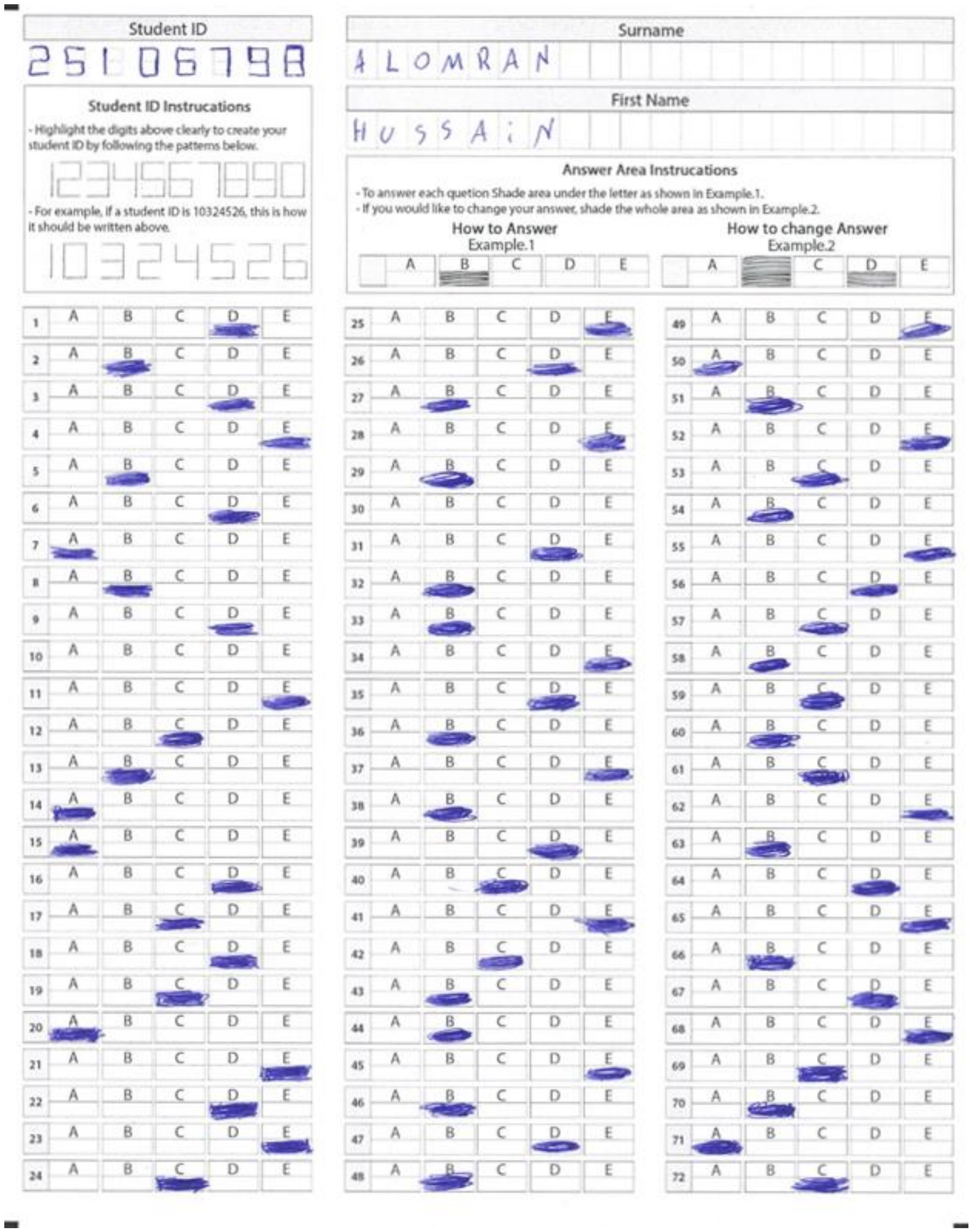


Figure 4.8: The cropped answer sheet after finder pattern detection.

Now, the answer sheet is ready to be processed. Hence, the marked answers will be recognised first. Figure 4.9 shows the segmented answer area. To detect the answers, every column of the answer area is segmented and then binarised as shown in Figure 4.10. As every answer box location is pre-calculated, every answer box is checked. If an answer box contains at least 60 pixels, this indicates that the answer box is shaded. If the answer box is shaded, the detected box is compared with the answer key. Later, based on the answer key, the score is determined for every answer box. Figure 4.11 shows the detected shaded boxes of the first answer area column, where the blue highlighted boxes are the correct answer based on the answer key, and the red highlighted boxes are the detected wrong answers based on the answer key.

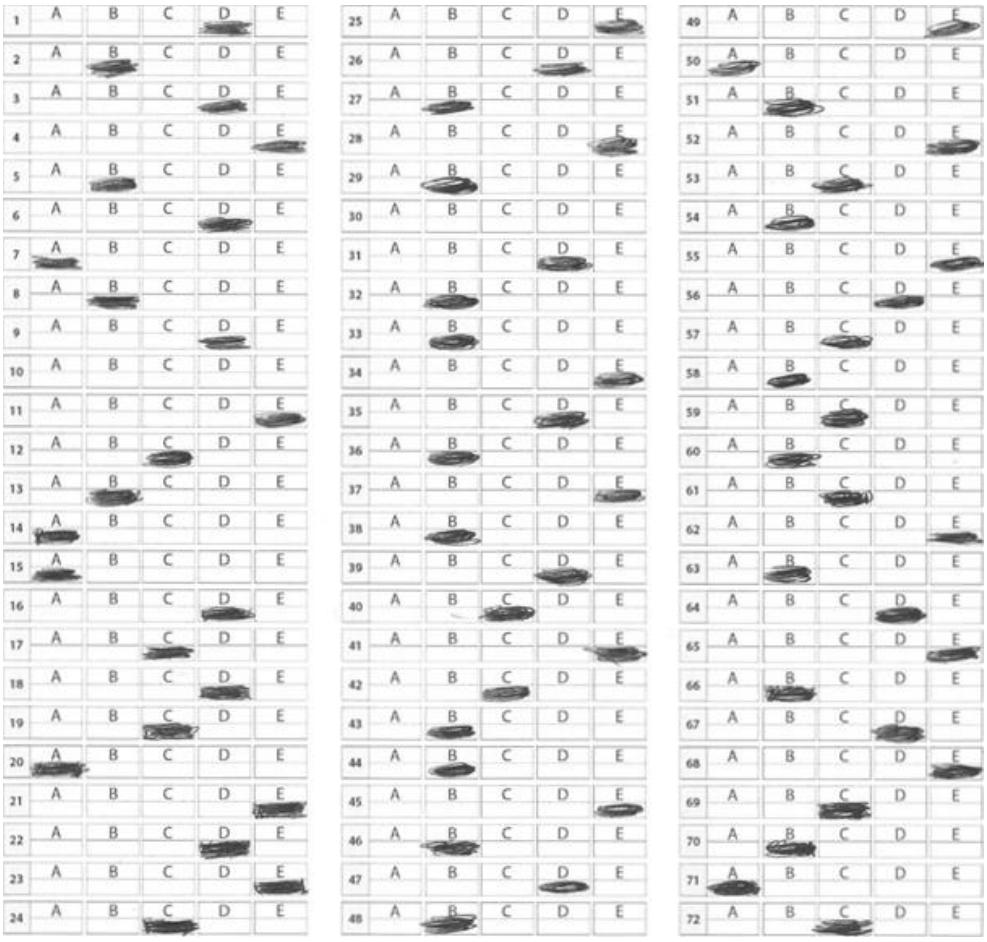
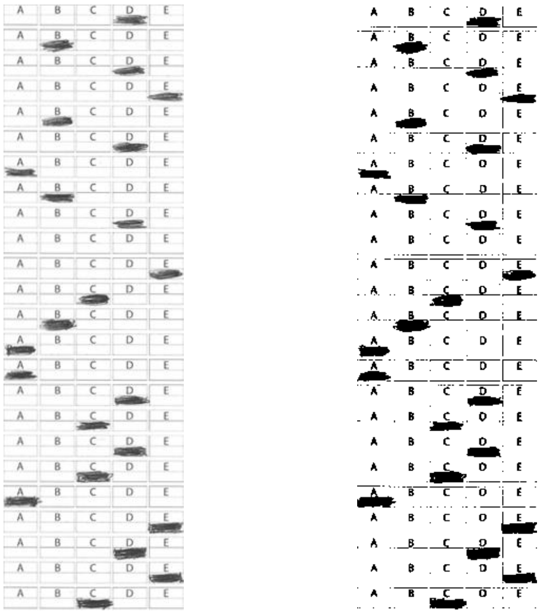


Figure 4.9: The segmented answer area.

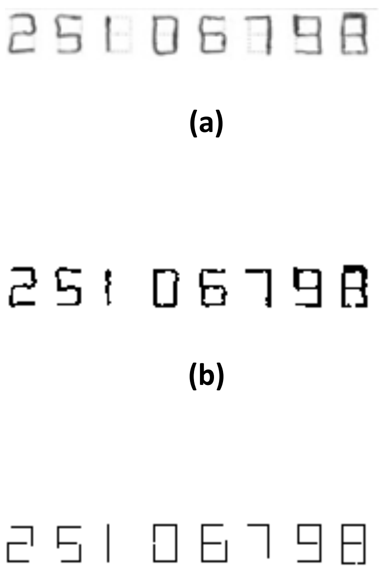


(a) **(b)**

Figure 4.10: (a) is the first answer column and (b) is the binarised column.

Figure 4.11: The scored answer column, where the blue highlighted boxes are the correct answer based on the answer key, and the red highlighted boxes are the detected wrong answers based on the answer key.

After obtaining the overall score of the answer area, it is the time to detect the student information. Student information section consists of student ID and name areas. The information that are needed from this section are the student ID and the surname character number. First, student ID area is segmented to be recognised as shown in Figure 4.12 (a). Next, the area is converted into binary image as shown in Figure 4.12 (b). As it is clear in Figure 4.12 (c) that the ID digits have lots of noise, a noise reduction method is used to smooth and thin the digits. This is achieved via taking the skeleton of each digit and removing the small disconnected pixels. The result of the noise reduction is shown in Figure 4.12 (d). Form Figure 4.12, it can be deduced that the thickness of the pen stroke is that matter as long the ID written on the digit segments properly. Next, the segments of every digit are checked with a threshold. If a segment has at least 6 pixels, then is registered as an active segment. After registering all the digit segments, the segments are compared with the active segments table (see Table 1). Then, the active segments are multiplexed into a number based on the active segments table. The final decoded ID is 25106798.



(a)

(b)

(c)

Figure 4.13: The stages of student ID noise reduction.

After decoding the ID, surname character count is checked. This aim is to speed up student information matching by skipping the ones with different surname character count. To count the surname character, first, surname area is cropped as shown in Figure 4.13 (a). Then the image is binarised as shown in Figure 4.13 (b). Every character position is pre-calculated and segmented on the answer sheet. the ID is matched with existing student database. Now basically, every pixel group of surname character is counted. If the pixel count meets the threshold, the algorithm counts up. As it can be seen in the figure that the surname consists of seven characters. Hence, the student who has an ID 25106798 with the only seven surname characters is searched in the database. If the student information is found, the overall test score will be stored with a high confidence flag. If not, the ID will be checked for auto-correction as discussed in Section 3.2.2. Note that in this example student information of this example were matched correctly with a high confidence.

(a)

(b)

Figure 4.13: Surname character processing.

The last part for the answer sheet to be processed answer sheet annotation and result feedback. Regarding answer sheet annotation, the location of every detected marked answer is stored. Based on that, the answer sheet is annotated with colouring style as shown in Figure 3.19. lastly, the annotated answer sheet is sent to the recognised student email from the database.

4.4 Algorithm Accuracy

In this section, the accuracy of the software algorithm will be investigated. First, we will be looking for the maximum rotation angle can the answer sheet be corrected. In addition how robust the finder pattern recognition algorithm is to noise. Furthermore, more examples of student IDs testing will be examined. In addition, several shading style will be investigated which shows how accurate the marked answer recognition algorithm is.

4.4.1 Finder Pattern

As it is shown in Section 4.3, the finder pattern points can be recognised. A successful finder pattern recognition means that the answer sheet is ready to be scored. However, an unsuccessful finder pattern recognition means that the software algorithm cannot proceed processing. Therefore, the factors that limits or affects finder pattern recognition must be examined. The most common factors that negatively affect finder pattern recognition are:

- Tilt during scanning answer sheets
- Distortion exhibited by a scanner
- Noise around finder pattern points.

During scanning, a small degree of rotation must exist as shown in Figure 4.4 which has a 0.32-degree tilt. The finder pattern recognition algorithm is designed to correct the inevitable tilt. Several experiments were conducted to test the maximum tilt degree that can be corrected via the algorithm. It was deduced that the maximum tilt degree can be corrected is 5-degree. Figure 4.14 shows a scanned answer sheet with 5-degree of deliberate tilt. After running the algorithm script, the answer sheet was successfully recognised and corrected. Figure 4.15 shows the answer sheet after rotation correction. As it can be seen from Figure 4.14, the answer sheet can be rotated up to 5 degrees. However, more than 5-degree tilt can obviously crop the finder pattern points. Hence, the algorithm will not be able to identify the points.

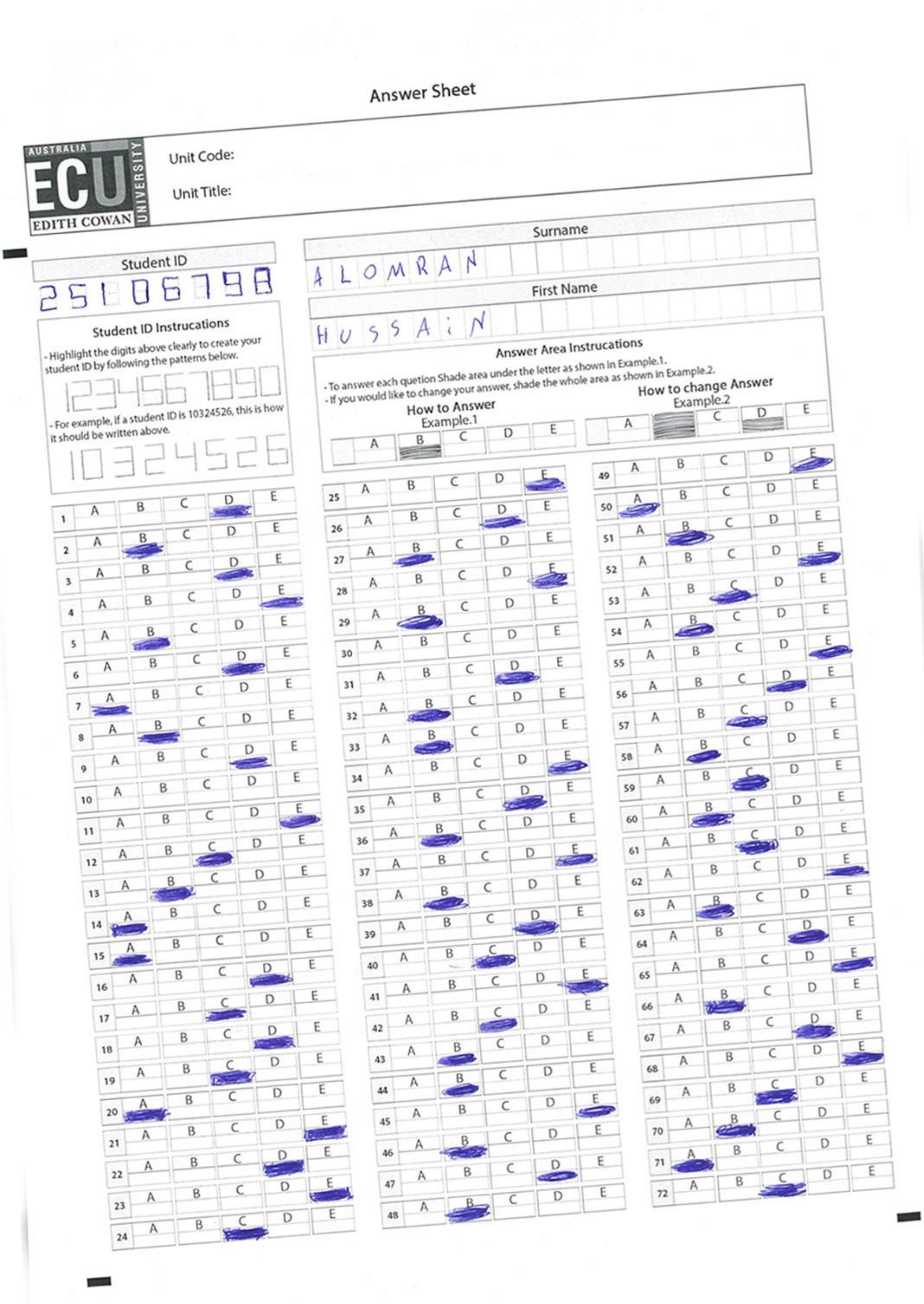


Figure 4.14: A scanned answer sheet with 5-degree tilt.

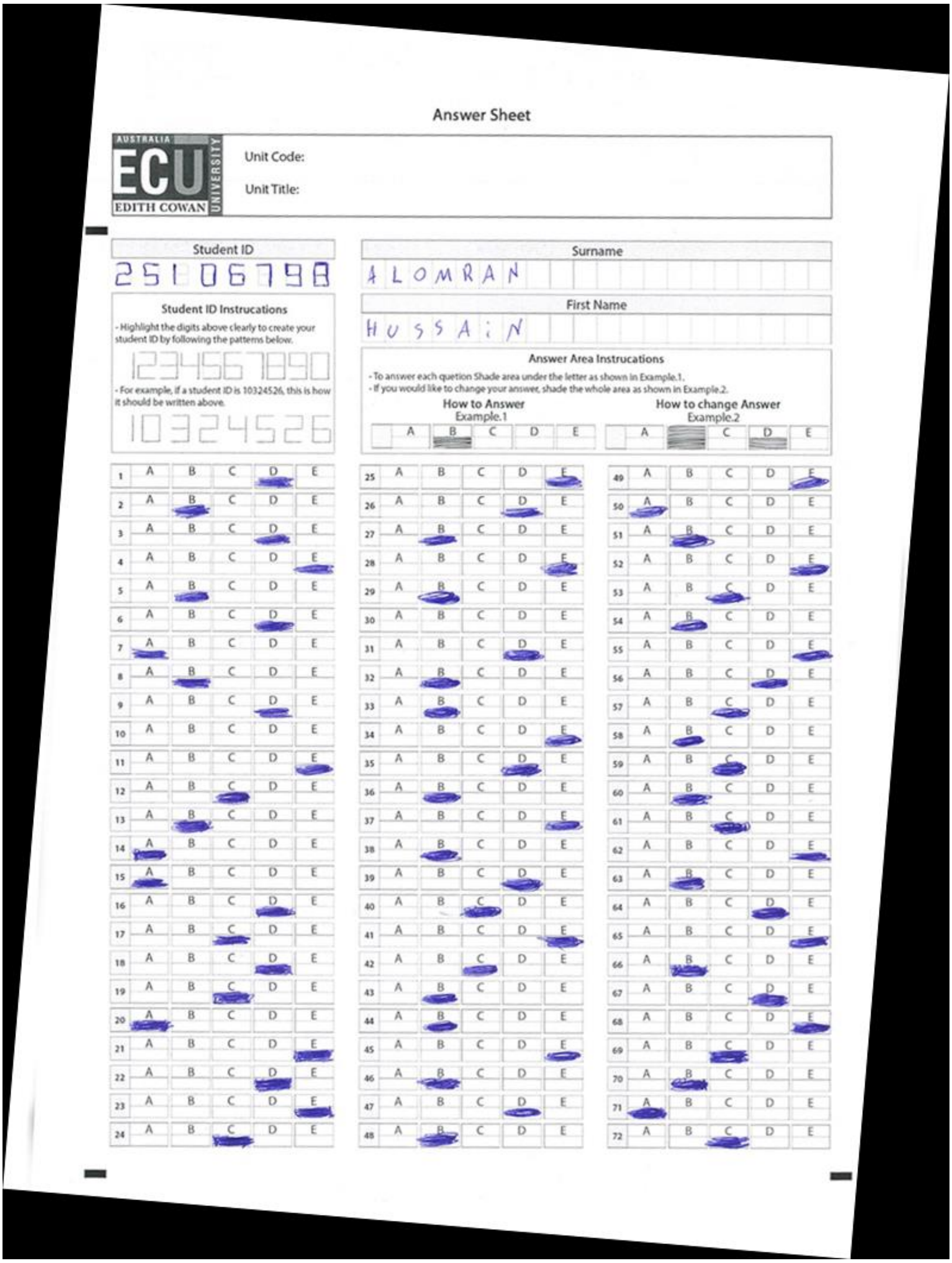


Figure 4.15: The corrected answer sheet by 5-degree in the opposite direction.

As shown in Figure 4.14, the answer sheet is free from distortion in the edges of the answer sheet. However, some cheap scanners exhibit dark spots in the frame. Two of shelf scanners were used to experiment the algorithm robustness to distortion in the edges. The first example is shown in Figure 4.16. The answer sheet in Figure 4.16 is scanned with small dark spots in the right and left margins. The answer sheet has 2.5-degree anticlockwise tilt too. The answer sheet was able to recognise the finder pattern and correct the sheet to the right alignment. The result is shown in Figure 4.17.

A second scanner that exhibits distortion in the margins was test. The result of the scanning is shown in Figure 4.18. As it can be seen for the figure, the answer sheet is severely distorted in the edge of the image frame. Moreover, the answer sheet was scanned with 3-degree tilt. The algorithm was able to erode the noise and locate the finder pattern points. The result of noise segmentation is shown in Figure 4.19. The figure shows the segmented areas that are bigger than the finder pattern points which are the logo and the dark spots in the margin. Note that the top right noise was eroded after binarisation process. The final result of the image tilt correction is shown in Figure 4.20.

Finally, the same answer sheet was scanned with same scanner but with pen scribbles near the finder pattern to test whether the algorithm will be confused with pen scribbles or not. The scanned answer sheet with pen scribble is shown in Figure 4.21. The result in Figure 4.22 shows that pen scribbles does not affect finder pattern recognition unless the scribbles touch one of the finder pattern points.

In conclusion, finder pattern recognition algorithm is able to recognise the finder pattern with tilt during scanning, distortion exhibited by a cheap scanner, and random pen scribbles near finder pattern points. However, the algorithm can correct the tilt up to 5-degree in both directions clockwise or anticlockwise. The algorithm can handle scanner distortion and no matter the size of the dark spots unless the distortion do not cover the finder pattern points. Lastly, pen scribbles do not limit the algorithm to recognise the finder pattern points if the scribbles do not cross the finder pattern points.

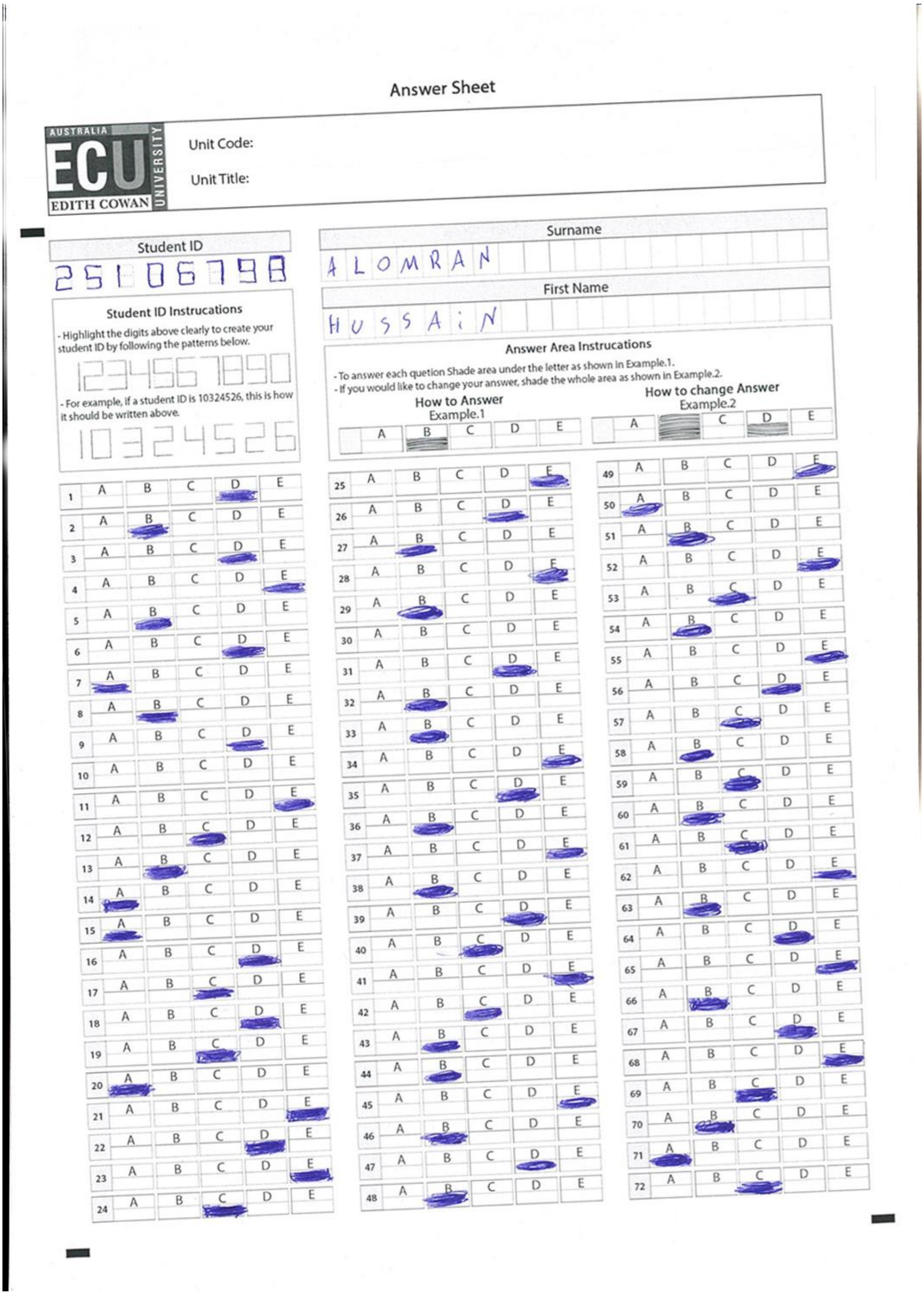


Figure 4.16: A scanned answer sheet by an off shelf scanner with 2.5-degree tilt.

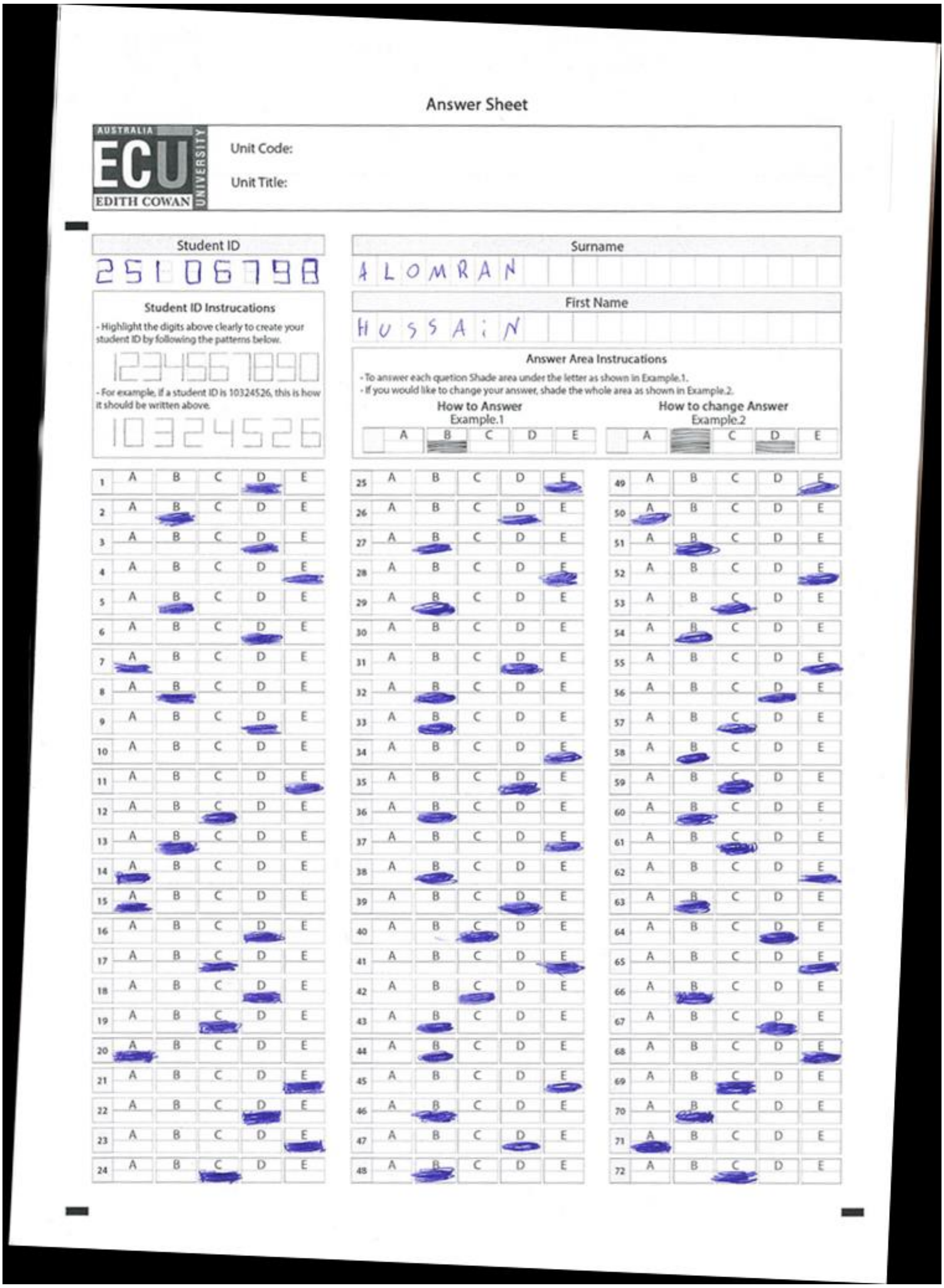


Figure 4.17: The corrected answer sheet.

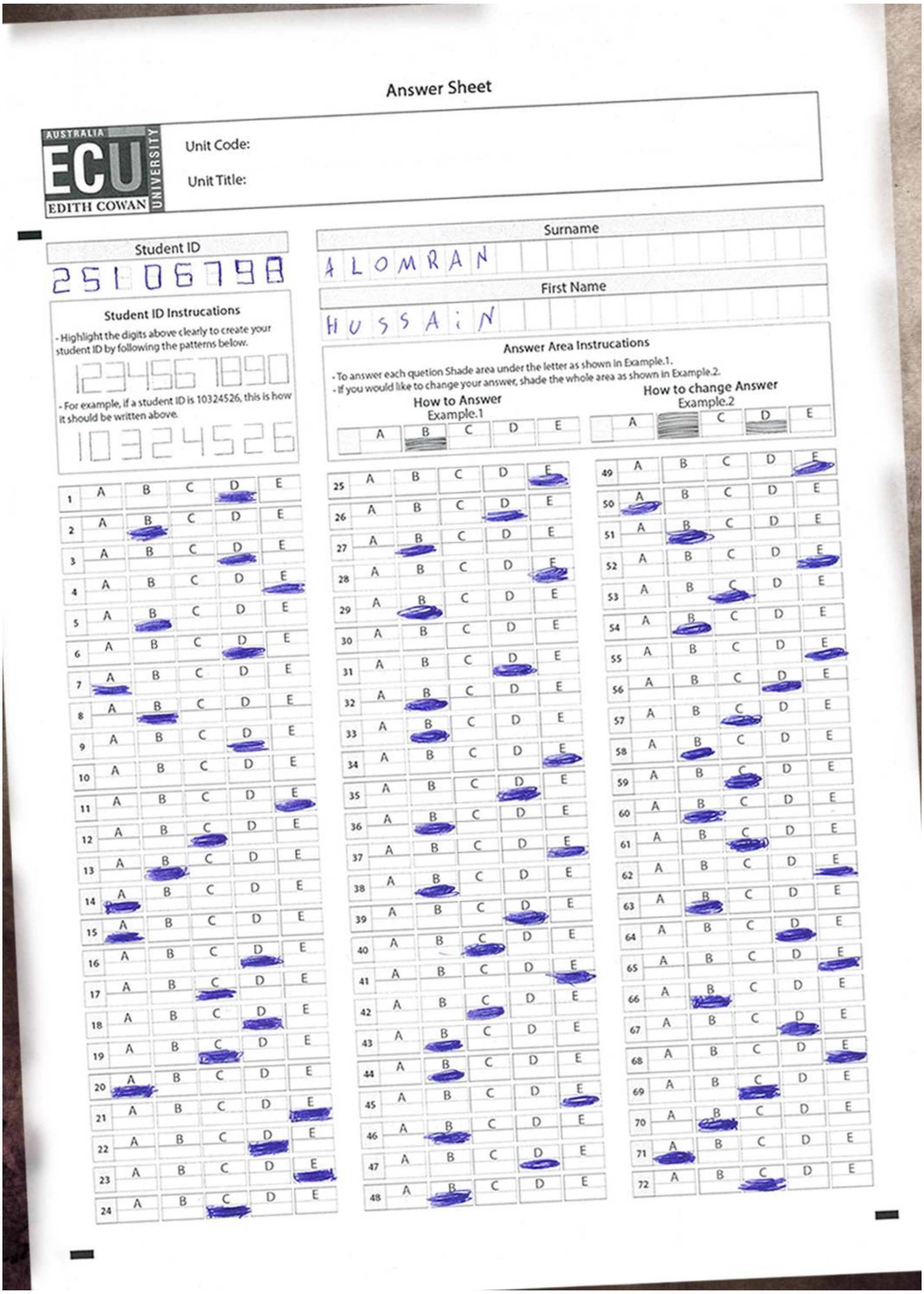


Figure 18: A Scanned answer sheet with severe distortion in frame edge with 3-degree tilt.

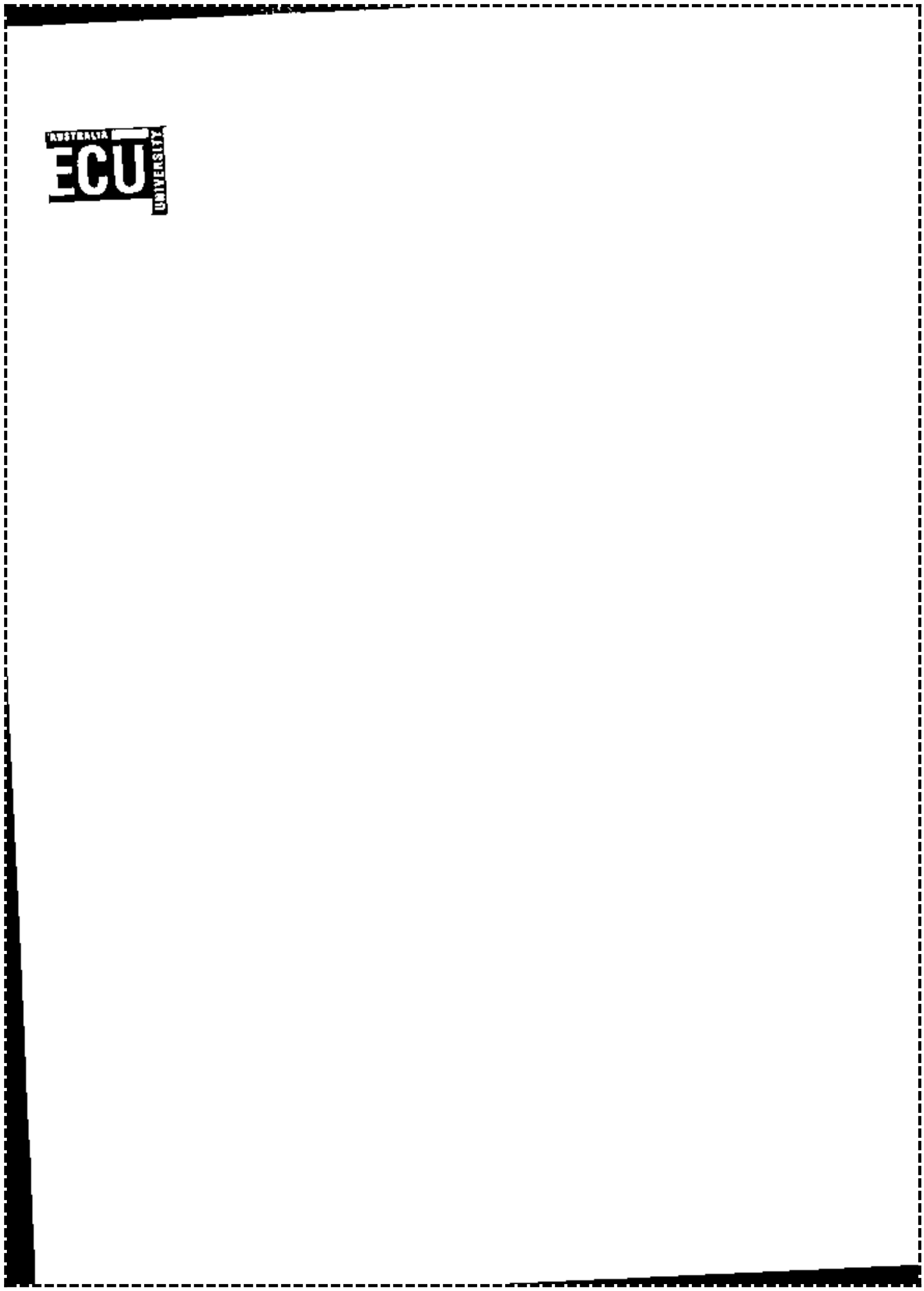


Figure 4.19: The segmented noise that is bigger than the finder pattern point.

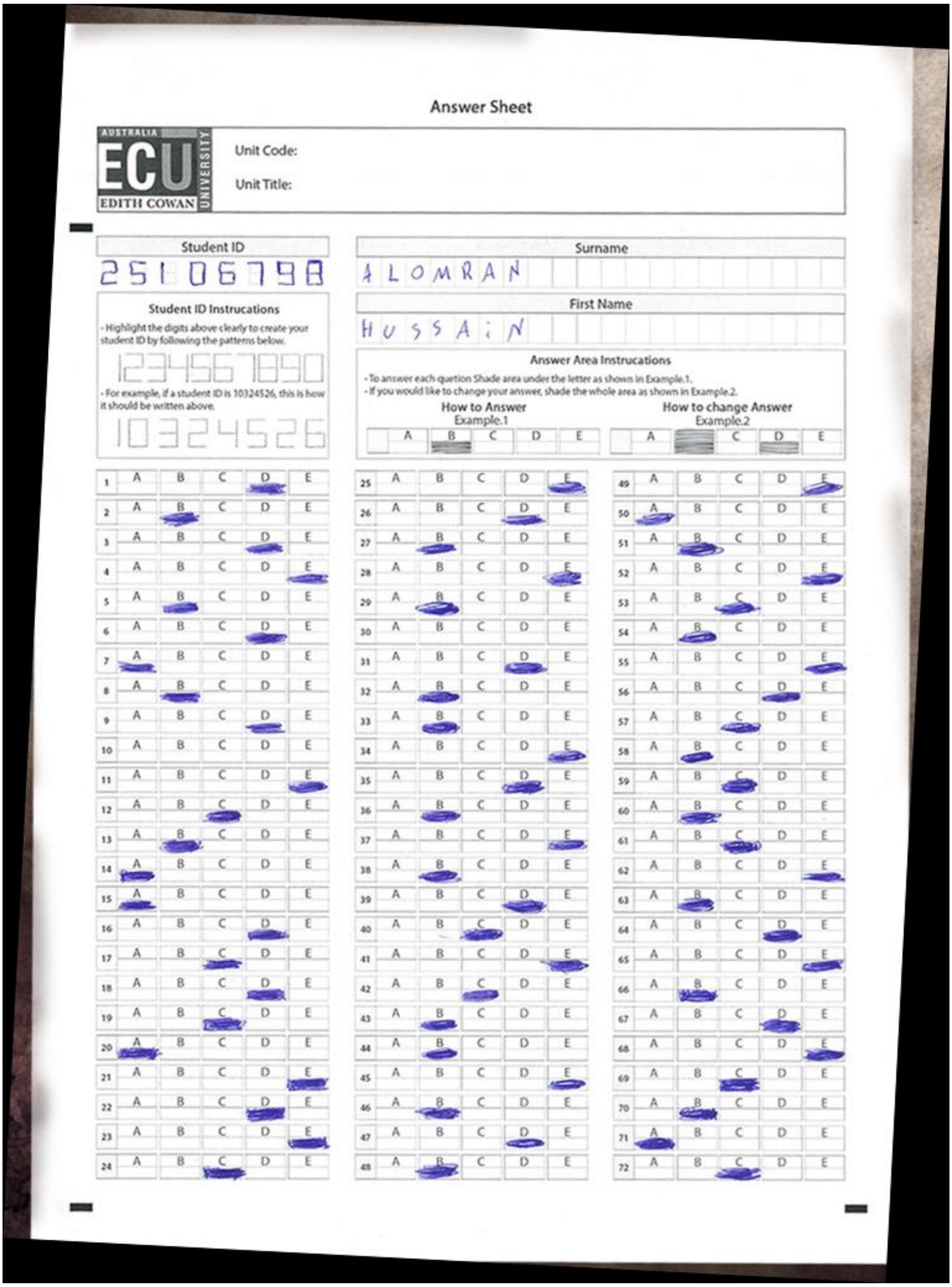


Figure 4.20: answer sheet tilt correction of 3-degree tilt and severe margin distortion.

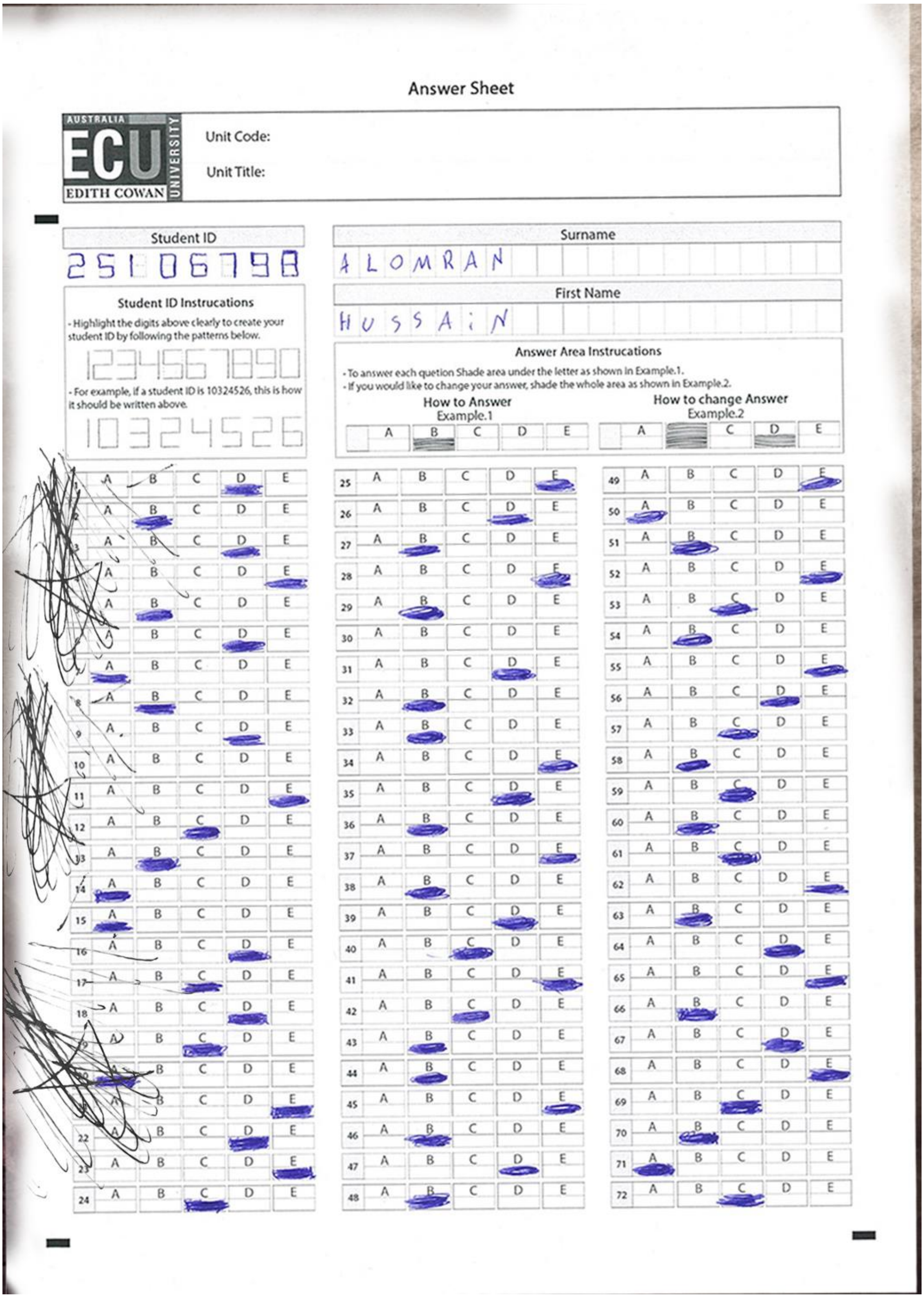


Figure 4.21: A scanned answer sheet with pen scribbles.

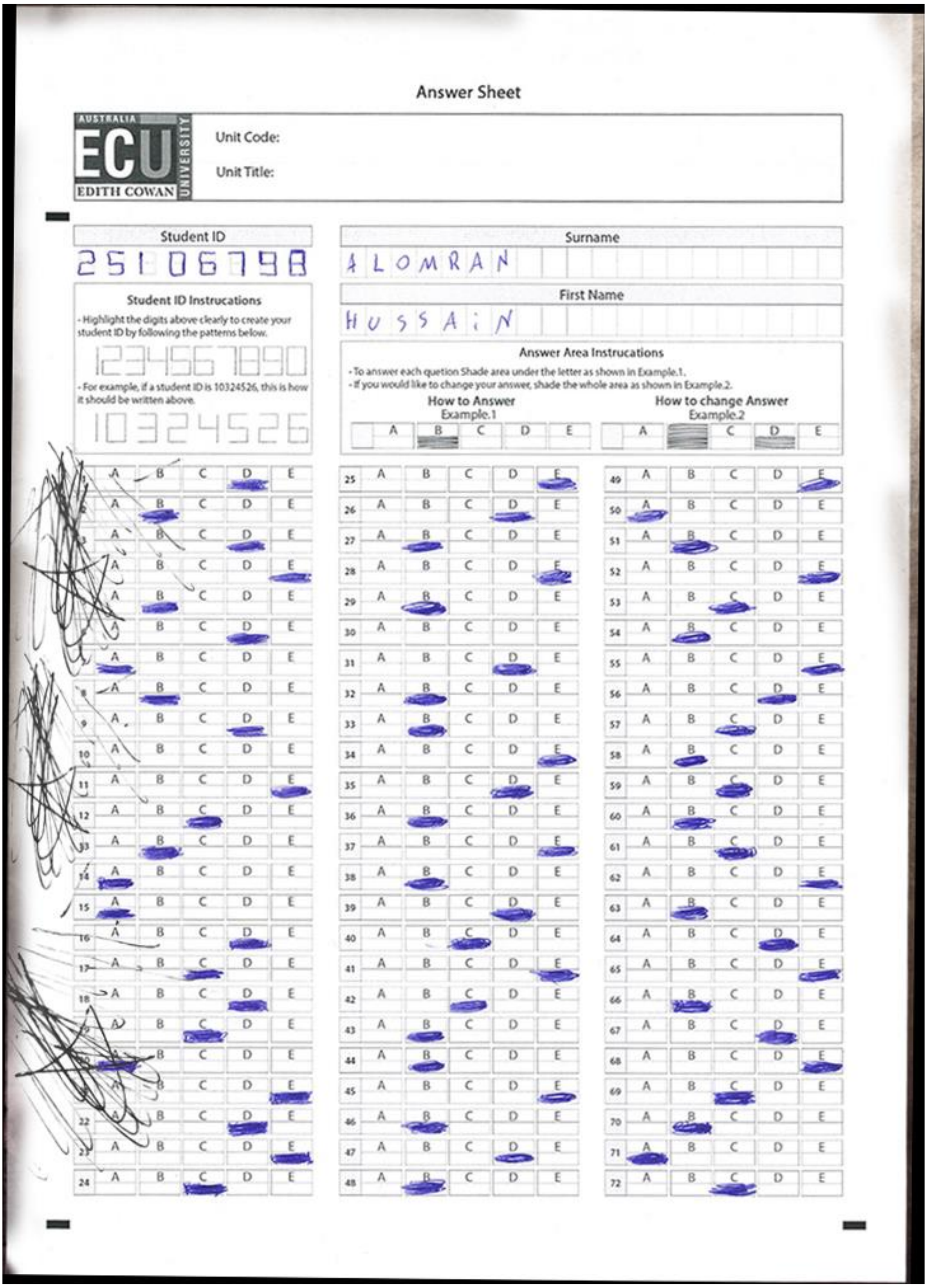


Figure 4.22: Corrected answer sheet tilt with pen scribbles near finder pattern area.

4.4.2 Student ID Recognition

Recognising student ID is the most importing part after finder pattern recognition in software algorithm. Because in this research we put more emphasis on the novel student ID recognition, the results of the novel student ID recognition method are important to study. The experimenting method was done in two parts. This first one is that every number was tested with different styles individually. This aims to study every number and the possible ways to decode it. Then, collective numbers of student IDs were tested.

The student ID recognition algorithm was tested with various styles. The algorithm proved to be accurate after testing over 100 numbers individually. To show a result, numbers from zero to nine were experimented as shown in Figure 4.23. Every number was written eight times with different styles. Under the handwritten numbers are the recognized numbers. Furthermore, number nine and six can be written in two different style as shown in Figure 4.23. Number one can be recognized correctly even if it is written in any side of the segment. This is because with number one, the width of the character is checked first, if the width size is less than the segment size, the number is recognized as one. As shown in Figure 4.23, if the numbers are written on the proper segments, the proposed algorithm will decode the number correctly with a 100% accuracy rate.

Many random student IDs were tested to evaluate student ID recognition accuracy. Figure 4.24 shows some examples of the recognised IDs. As shown in the figure, the thickness of the pen is not important as long the pen stroke on the right segments. Furthermore, the algorithm was test if it can recognise a student ID with noise in between the active segments as shown in Figure 4.25. The algorithm was successfully able to decode the ID. This is because that the algorithm checks only the active segments and ignores any pixel outside the segments. However, the algorithm was able to recognise an ID was written slightly outside the active segments which is shown in Figure 4.26. As shown in the figure, number 2, 4, and 5 were written slightly above the grid. Hence, these number were not recognised. In addition, an encoded ID with handwritten style was tried out which is shown in Figure 4.27. The algorithm was not able to recognise the numbers.

In conclusion, the student ID recognition algorithm is able to decode student IDs with different pen thickness as long the numbers are written in the grid properly. Furthermore, the algorithm proved to be robust to noise around the active segments. However, the algorithm is capable to decode any number written with handwritten style or outside the grid. All the numbers have to be written as the style of the grid to allow the algorithm to recognise the active segments.

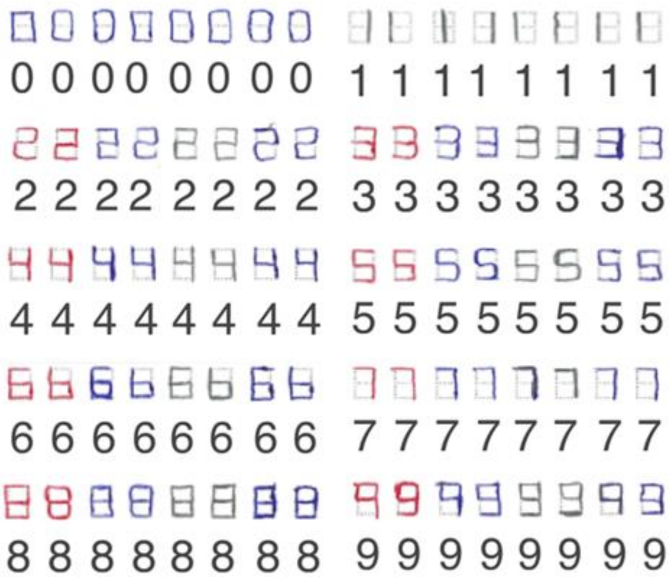


Figure 4.23: Character recognition test of the ten numbers.

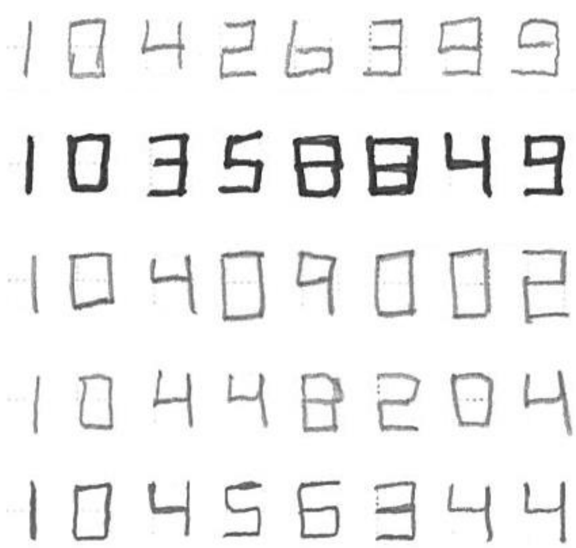


Figure 4.24: Some examples of the tested IDs.

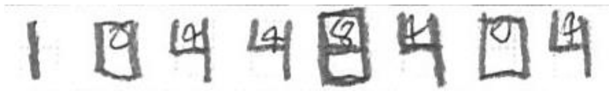


Figure 4.25: An ID recognition with noise in-between the active segments.

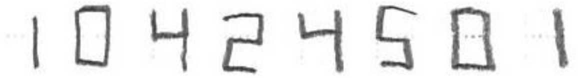


Figure 4.26: An ID test with some numbers outside the grid.

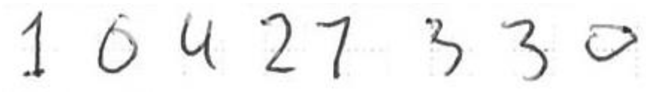


Figure 4.27: An ID test with handwritten style.

4.4.3 Marked Answer Box

The testing of the marked answer detection consists of two parts. First, the answer area was experimented with different shading styles and intensities. Then, the robustness of the marked answer recognition was tested with noise in the answer area.

As the marked answer recognition algorithm was implemented based on the four possible cases for each answer item: no entry, single entry, multiple entries, and single entry with changing answer attempt as shown in Figure 3.17, the four cases were tested. Over thousand marked answer with same style as shown in Figure 3.17 were detected with 100% success rate. Figure 2.49 shows two examples of marked answer recognition testing. As shown in Figure 4.29 (a), Q1 has no entry, hence no mark was detected. Furthermore, Q2 and Q6 have a single entry and it can be seen that the marks are highlighted with a red rectangle. Q3 and Q5 show a single entry with changing answer attempts. The marked answers are ignored because change answer areas are shaded and the marked area without changing answer attempt is registered. Moreover, the algorithm is able to recognize if there are multiple entries as shown in Q4, hence both marks are not scored. Figure 2.49 (b) shows another samples of marked answer recognition. It can be seen in Q6 that any style can be recognized as long as the shape covers at least 40% of the answer box which is the recognition threshold.

However, all the samples that are shown in Figure 4.28 were tested with similar pen intensities. After several testing, it can be deduced that the algorithm can recognise ink pen robustly. In addition, the algorithm proved that it is robust as well with light pencil intensity. An example of light pencil shading is shown in Figure 4.29. Furthermore, more tests were carried out on the answer area under noise. The most common noise that the answer area that may encounter is some pen scribbles while shading the answer boxes. Figure 4.30 (a) shows an example of pen scribbles inside the answer area. It can be seen that the algorithm was able to score the marked answer boxes did not have an issue with the scribbles.

In conclusion, the marked answer algorithm is able to detect and process the answer boxes. Moreover, the algorithm is able to recognise the four possible cases such as no entry or multiple entries etc. Furthermore, the algorithm can detect different shading styles if only they cover at 40% of the answer box. In addition, the algorithm proved that it is robust with different pen intensities such as ink pen or light pencil intensity. Finally, the algorithm is tolerant with pen scribbles in the answer area.

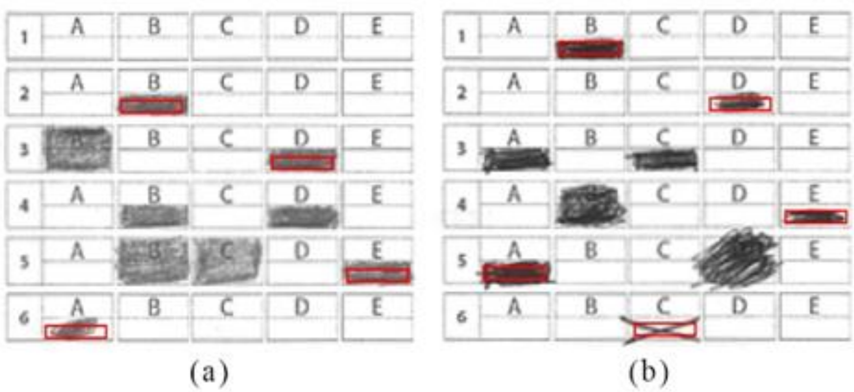
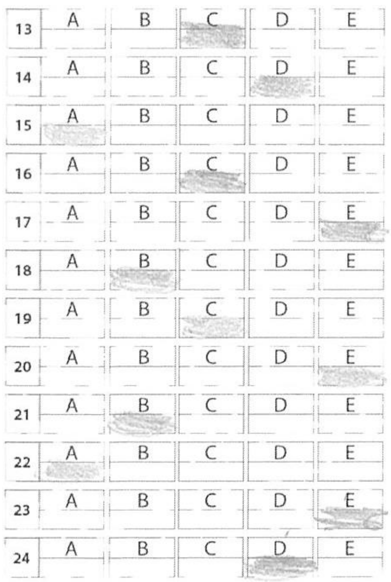
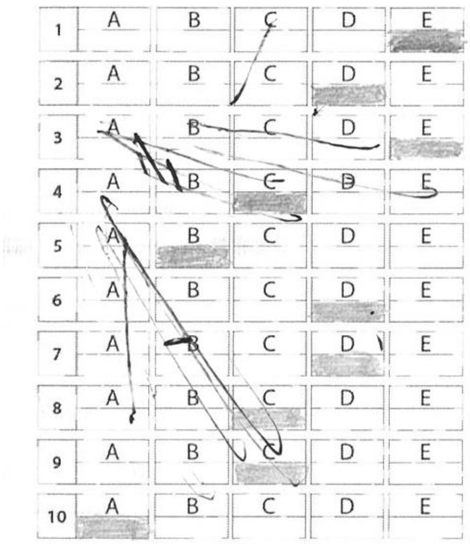


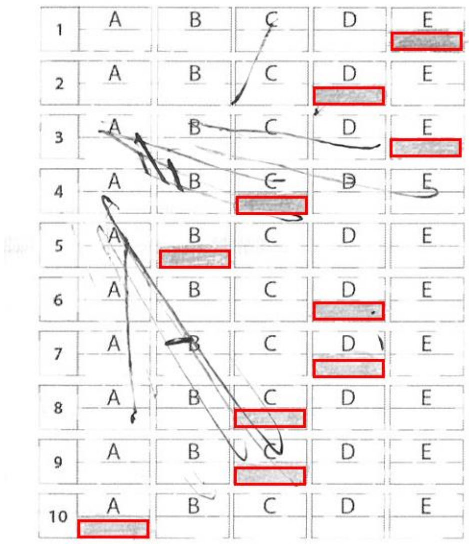
Figure 4.28: An example of marked answer recognition testing.



(a) **(b)**

Figure 4.29: An example of light pencil shading detection, (a) is the actual answer area, and (b) is the answer area after detection.





(a) **(b)**

Figure 4.30: An example of marked answer detection with pen scribbles in the answer area, (a) is the actual answer area, and (b) is the answer area after detection.

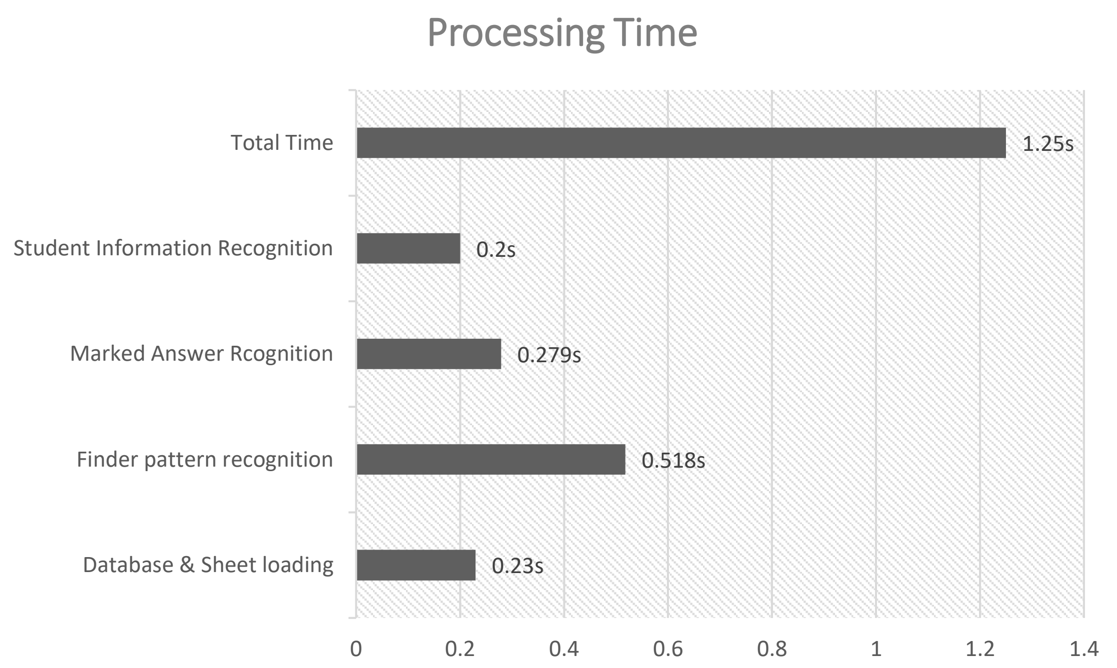
4.5 Algorithm Speed

As it is aimed to implement an accurate scoring algorithm, the algorithm processing time is important as well to investigate.

A group of answer sheets were processed automatically in an aim to calculate the processing time. The tool was used to calculate the processing time is Profiler which is a built-in function in MATLAB. The Profiler function offers the able to calculate the overall processing time and the time of each script command. Therefore, the overall processing time will be discussed and quickest and the slowest in time of the algorithm script will be analysed. Furthermore, a comparison between the algorithm speed and one of the scoring machines will be covered.

The processing time of scoring 100 answer sheets is 125 seconds. This is equivalent to 1.25 seconds per answer sheet. This time includes scoring an answer sheet without annotation. The processing time of scoring and annotating of 100 answer sheets is 227 seconds which is equivalent to 2.27 seconds per answer sheet. This is the time from reading the image from the hard drive to saving the annotated sheet. The processing time of an annotated sheet is longer than the processing time without annotation. This is because after annotating the sheets, every annotated will be saved in the storage directory. Saving an annotated sheet takes almost 1 second which explain why the processing time is longer than the time without annotation. The two commands that take most of the algorithm time are reading the answer sheets from the storage directory and saving them according to the Profiler. Figure 4.31 illustrates the time of the algorithm parts. From the figure, the processing of finder pattern recognition takes the longest part of algorithm in time. Furthermore, the shortest part of the algorithm to process is student information recognition Note that the processing time may vary based on the computer RAM size and processor speed.

The processing time of the system is much better than manual marking. It takes almost 7 minutes to score the same proposed answer sheet and allocate the score for each student on a spreadsheet. The average processing time of a fast optical mark machine to score an answer sheet is almost 1.4 seconds [77]. For example, if we compare Scantron OpScan 8 which costs around US$11335 and is one of the most popular scoring machine, the proposed system is faster and much cheaper. Table 3 shows the comparison between manual scoring, Scantron OpScan 8, and the proposed system.



	Database & Sheet loading	Finder pattern recognition	Marked Answer Rcognition	Student Information Recognition	Total Time
Processing Time in Seconds	0.23	0.518	0.279	0.2	1.25

Figure 4.31: Processing time of the algorithm parts.

Table 3: Scoring Methods comparison in time.

Manual Scoring	OpScan8	Proposed System
7 Minutes	1.39 seconds	1.25 second

4.6 Case Study

After obtaining the research ethic approval for the Research Ethic Office at Edith Cowan University, the proposed system was conducted in a multiple-choice class test. The total number of the case study participants is 183 students. All participants had to use the proposed answer sheet to answer the test questions. The aim of the case study is to understand student response to answer sheet design and examine the robustness of the proposed scoring algorithm. Hence, the finder patterns, marked answers, and student information recognition algorithms were the main results to analyse. Note that students were not given additional instructions on how to fill out the answer sheet. The only given instructions were on the answer sheet.

After finishing the test, the 183 answer sheets were scanned via Ricoh MP C5501A. Then, after providing the answer key and the class list (database), the algorithm processed all the answer sheets and generated the Excel sheet in almost 90 seconds.

In all the answer sheets, the algorithm was able to recognise the finder pattern and scored 100% success rate in finder pattern recognition. Moreover, the success rate for marked answer recognition was 98.9%. This is because two participants used pencil and erased their answers but not properly as shown in Figure 4.32. Hence, the algorithm picked up some multiple entries as the algorithm is very accurate with light pencil intensities.

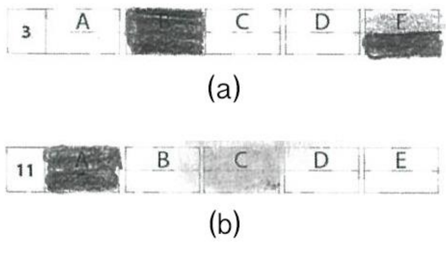


Figure 4.32: Two marks were not recognised due to shading intensity issues.

In addition, 160 student information out 183 were recognised correctly and matched with the class list. The reason that the 23 student information were not recognised is because their IDs were either written outside the segment grids or written in handwritten style. Figure 4.33 shows some examples of the unrecognised student IDs.

As it can be seen in Figure 4.33 (a) and (c), the IDs were written in a hand written style. Where in Figure 4.33 (b), the ID was written in the correct style, but two digits were crossed which led to unsuccessful ID decoding. Furthermore in Figure 4.33 (d), the ID was written and then with pen scribble which stopped the algorithm to recognise any digit.

Furthermore, 137 student IDs were matched with high confidence. 23 student IDs were matched with low confidence, which means that the auto-correction algorithm were able to this mis-decoded digits and correct them based on the student IDs list. Overall, the succes rate of the student information recognition was 90.9%.

Figure 4.34 demonstrates an example of a fully recognised answer sheet. The figure is annotated to show the decoded student ID, total score, and recognised marked answers by the scoring system. As shown in the figure, the recognised ID is annotated by the algorithm below the student ID area. Furthermore, the marked answers are annotated with different colours (red, blue, and green), where the red mark denotes a wrong answer based on the answer key, the blue mark is the corrected answer, and the green mark is the correct answer. Moreover, the total mark is shown in the top right of the answer sheet.

In order to check the credibility of the algorithm results, all the answer sheets were scored manually first. After processing the sheets via the algorithm, the scores and student information were compared with the manually scored answer sheets. The overall result of the case study is illustrated in Figure 4.35.

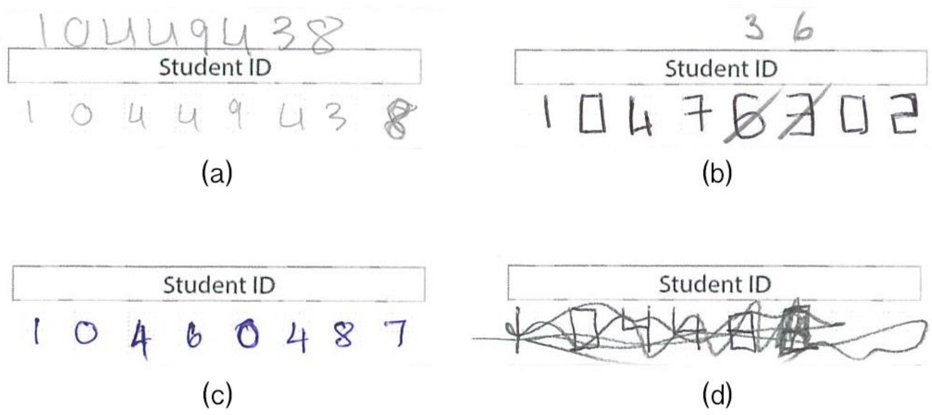


Figure 4.33: some of the unrecognised students IDs.

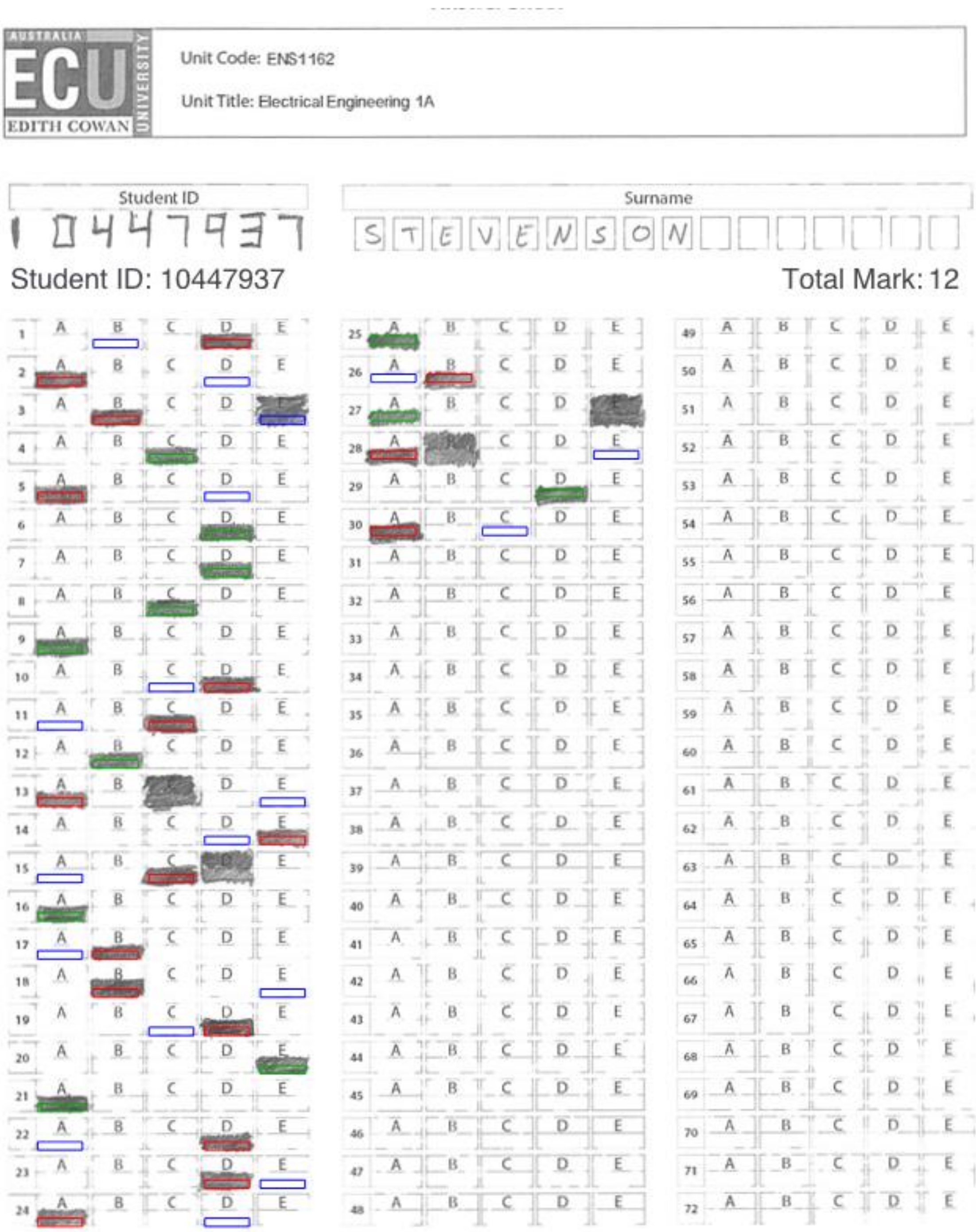


Figure 4.34: An example of a fully recognised answer sheet.

Figure 4.35: The overall case study result.

4.7 Results Summary

The summary of the implemented system results can be seen in Table 4.

Table 4: Results summary of the system.

Effectiveness	
Finder pattern recognition rate	100%
Mark recognition rate	98%
Changing answer recognition rate	100%
Student information recognition without auto correction rate	85%
Student information recognition with auto correction rate	91%
Robustness	
Can recognise answer sheet with different orientations (vertical/horizontal)	Yes
Maximum rotation degree answer sheet can be recognised	±3°
Can recognise answer sheet with distortion in the edges	Yes

Can recognise answer sheet with distortion around finder pattern	No
Can recognise answer sheet with distortion around answer area	Yes
Can recognise multiple answer correction	Yes
Can recognise multiple answer entries	Yes
Can recognise handwritten ID naturally	No
Speed	
Processing per answer sheet without annotation	1.25s
Processing per answer sheet with annotation	1.8s

4.8 Limitations

After conducting the individual experiments on the proposed system and the case study, the deduced limitations of the systems are as follow:

1. The algorithm can correct the scanning tilt of 5-degree maximum. More than 5-degree, the algorithm will not be able to recognise the finder pattern points. Moreover, the algorithm cannot recognise the finder pattern if the three points or one of them is distorted.
2. The algorithm cannot decode student IDs that are written with handwritten style. To ensure that the student information is recognised correctly, most importantly, the student ID has to be written as mentioned on the answer sheet instructions. The algorithm may be able to correct one or two of the ID digits, but decoding an ID without following the instructions may lead recognition with low confidence or no recognition at all.
3. The algorithm can detect marked answer with different colours and intensities. However, the algorithm may get confused with erased answer if the answer were not erased properly. Furthermore, the algorithm is able to detect any style of shading in the answer box. However, the shading area has to cover at least 40% of the answer box area.

4. All the scanned answer sheets have to be converted into JPEG image format to process them. The algorithm will not be able to process PDF files. Student database and answer key files have to be provided with Microsoft Excel format.
5. In order to utilise the automated scoring system, the designed answer sheet has to be used otherwise the algorithm will not be able to execute.

Overall, after obtaining and analysing the results of the proposed system, it can be claimed that the algorithm is fast with 1.25 seconds per sheet, reasonably accurate to overcome noise, and recognise student information and marked answer.

5. CONCLUDING REMARKS

5.1 Conclusion

In this research, an extensive literature review on assessment and automated multiple-choice scoring systems and explore the state of the art were reviewed. Furthermore, the design of a user friendly answer sheet were implemented. In addition to answer sheet design, an automated image processing-based scoring system were implemented and discussed. Moreover, the testing and robustness evaluation of the proposed system were carried out.

As marking multiple-choice test papers takes long time and the process of marking hundreds of test papers is not free of mistakes. Furthermore, the traditional marking method requires more human resources if there are many test papers to mark in a limited time, which is more expensive. The proposed system in this research is a fast and cost effective technique for automatic optical mark recognition.

The proposed system utilises image processing algorithms to decode student information and multiple-choice marks automatically. In addition, as part of the research contribution, the proposed system provides students their test papers with annotation via their emails. Moreover, not like the ordinary automated multiple-choice test scoring system, the proposed method allows students for multiple corrections to the answer sheet.

After experimenting the system, the system results indicates that the system meets the research objectives. Speed wise, the system scoring speed is 1.25 seconds and 2.25 seconds with annotation. The system speed has proved to be faster than one of the most common scoring machines as discussed in Chapter 4. Furthermore, the system was tested against distortions. The software algorithm proved to be robust to detect the finder pattern points when noise exists in the answer sheet margin and pen scribbles around the finder pattern area. In addition, the algorithm is able to correct scanning tilt up to 5-degree rotation. In addition, the novel student ID recognition algorithm was experimented to verify its robustness. The algorithm has 100% success rate when the individual testing was conducted. Furthermore, the marked answer algorithm was experimented to test the four possible cases (no entry, single entry, multiple entries, and single entry with changing answer attempt). The algorithm was able to recognise all the four cases. Also, the algorithm is capable to recognise different shading styles as long the shading area covers at least 40% of the answer box. A case study was conducted in a real test situation to retrieve more result about the system. The outcomes of the case study are that the succession rate of finder pattern recognition was 100%, the succession rate of marked answer recognition was 97.7%%, and the succession rate of student information recognition was 90.2%.

However, there are several system limitations. In order to utilise the automated scoring system, the designed answer sheet has to be used otherwise the algorithm will not be able to execute. Moreover, all the scanned answer sheets have to be converted into JPEG image format to process them. The algorithm will not be able to process PDF files. Student database and answer key files have to be provided with Microsoft Excel format. Furthermore, noise like pen scribbles on the finder pattern points will stop the algorithm to register the answer sheet which means that the execution of scoring will not be achieved.

5.2 Future Work

There are many aspects of the proposed system to improve such as implementing a handwritten optical character recognition to decode student IDs and names from the answer sheet. However, one of the main future work of the research is to design an unstructured answer sheet recognition. The proposed software algorithm is structured to recognise the designed answer sheet. Without the designed answer sheet, the proposed system cannot be utilised. Therefore, designing unstructured algorithm that recognises any answer sheet design will have a significant reduction on usability restriction. Furthermore, one of the future work in this research area is to design an online scoring system opposing to the proposed system which is an offline system. Providing the same proposed system but online will also reduce the restriction on the assessor to perform scoring anywhere.

REFERENCES

[1] D. Johnson, “Stretching Your Technology Dollar,” *Education Leadership,* vol. 69, no. 4, pp. 30-33, 2012.

[2] Carnegie Mellon University, “Assessment Basics,” Eberly colleague, 2015. [Online]. Available: https://www.cmu.edu/teaching/assessment/basics/index.html.

[3] S. Ramesh and M. Sidhu, “Exploring the Potential of Multiple Choice Questions in Computer-Based Assessment of Student Learning,” *Malaysian Online Journal of Instructional Technology,* vol. 2, no. 1, 2005.

[4] T. Roberts, “The Use of Multiple Choice Tests for Formative and Summative,” *Australian Computer Society,* vol. 52, pp. 175-180, 2006.

[5] Pearson Education, “How digital are Australian Schools?,” Pearson Education, [Online]. Available: www.pearson.com.au/insights-and-news/the-future-of-education/how-digital-are-australian-schools/. [Accessed March 2018].

[6] J. Noyes and K. Garland, “Computer- vs. paper-based tasks: Are they equivalent?,” *Ergonomics,* vol. 51, no. 9, pp. 1352-1375, Sep 2008.

[7] J. Fisteus, A. Pardo and N. Garcıa, “Grading Multiple Choice Exams with Low-Cost and Portable,” *Springer Science+Business Media,,* vol. 22, no. 4, pp. 560-571, 2012.

[8] V. Clegg and W. Cashin, “Improving Multiple-Choice Tests,” *iDEA PAPER NO.16,* pp. 208-2012, 1986.

[9] W. Harlen and M. James, “Assessment and Learning: differences and relationships between formative and summative,” *Assessment in Education,* vol. 4, no. 3, pp. 365-379, 28 Jul 2006.

[10] D. Wiliam, “Assessment: The Bridge between Teaching and Learning,” *Voices from the Middle,* vol. 21, no. 2, pp. 15-20, Dec 2013.

[11] ACT Government, “Teachers’ Guide to Assessment,” June 2011. [Online]. Available: https://www.education.act.gov.au/__data/assets/pdf_file/0011/297182/Teachers_Guide_to_Assessment_Web.pdf. [Accessed 2018].

[12] Employment and Workplace Relations for the Council of Australian Governments, “The Early Years Learning Framework for Australia,” the Australian Government , Barton , 2009.

[13] E. Wood, “What are Extended Matching Sets Questions?,” *Bioscience Education,* vol. 1, no. 1, pp. 1-8, 2003.

[14] Hanover Research Council, “Student Demand for Alternative Modes of Course Delivery,” Hanover Research Council, Washington DC, 2009.

[15] H. Wang , “Ranking Responses in Multiple-Choice Questions,” *Journal of Applied Statistics,* vol. 35, no. 4, pp. 465-474, Apr 2008.

[16] S. Merritt, Mastering Multiple Choice: The definitive guide to better grades on multiple choice exams, 6th ed., Canada: Brain Ranch, 2006.

[17] M. Kastner and B. Stangl, “Multiple Choice and Constructed Respons Tests: Do Test Format and Scoring Matter?,” in *International Conference on Education and Educational Psychology*, Vienna, 2011.

[18] T. Roberts, “The Use of Multiple Choice Tests for Formative and Summative Assessment,” *Informatics and Communication,* vol. 52, no. 2, pp. 175-180, 2006.

[19] V. Soudien, “Electronic Marking Literature Review,” *Computer Science,* vol. 2, no. 1, pp. 1-6, May 2014.

[20] IBM, “Automated Test Scoring,” International Business Machines , [Online]. Available: http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/testscore/.

[21] N. V. Patel and I. G. Prajapati, “Various Techniques for Assessment of OMR Sheets through Ordinary 2D Scanner: A Survey,” *International Journal of Engineering Research & Technology (IJERT),* vol. 4, no. 9, pp. 803-807, Sep 2015.

[22] International Business Machine, “Technical Breakthroughs,” IBM, [Online]. Available: http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/testscore/breakthroughs/. [Accessed 1 Sep 2016].

[23] International Business Machine Corporate, “Custumer Engineering Service Aids,” [Online]. Available: http://ed-thelen.org/comp-hist/IBM-ProdAnn/1230.pdf. [Accessed 01 Sep 2016].

[24] Optical Mark Recognition, “OMR Sheet Template,” [Online]. Available: http://optical-mark-recognition.com/omr-sheet-template.html. [Accessed Sep 2016].

[25] C. Higgins, “Pencils Down: Scantron Inventor Michael Sokolski Has Died,” Mental_Floss, 29 Jun 2012. [Online]. Available: http://mentalfloss.com/article/31057/pencils-down-scantron-inventor-michael-sokolski-has-died.

[26] Scantron Corporation, “Scantron: Our Story,” Scantron Corporation, 2016. [Online]. Available: http://www.scantron.com/about-us/company/our-story.

[27] Scantron Corporation, “EZData Scanner: Portable, accurate, and cost-effective,” Scantron , Eagan, 2014.

[28] Scantron Corporation, “iNSIGHT 4ES OMR and Image Scanner,” Scantron, Eagan, 2015.

[29] Scantron, “OpScan 4ES Scanner,” Scantron, Eagan, 2014.

[30] Scantron Corporation, “iNSIGHT 150 Scanner: Smarter imaging technology solution,” Scantron, Eagan, 2014.

[31] L. O’Gorman and R. Kasturi, Document Image Analysis: An Executive Briefing (IEEE Computer Society Executive Briefing), New York: Institute of Electrical & Electronics Enginee, 2009, pp. 1-125.

[32] P. S. Sandhu, G. Singla and S. Gupta, “A Generalized Approach To Optical Mark Recognition,” *International Conference on Computer and Communication Technologies (ICCCT'2012),* pp. 1-3, 27 May 2012.

[33] Remark Products Group, “Remark Office OMR,” Gravic, Malvern, 2012.

[34] K. Chinnasarn and Y. Rangsanseri, “An image-processing oriented optical mark reader,” *Society of PhotoOptical Instrumentation Engineers,* 1999.

[35] T. Nguyen, Q. Manh, P. Minh, L. Thanh and T. Hoang, “Efficient and reliable camera based multiple-choice test garding system,” *International Conference on Advance Technologies for Communications,* pp. 268-271, 2011.

[36] R. China, F. Zampirolli, R. Neves and J. Gonzalez, “An Application for Automatic Multiple-Choice Test Grading on Android,” *Revista Brasileira de Iniciaçã,* vol. 3, no. 2, pp. 4-25, 22 Feb 2015.

[37] F. Zampirolli , J. Gonzalez and R. Neves, “Automatic Correction of Multiple-Choice Tests using Digital Cameras and Image Processing,” *Universidade Federal do ABC,* 2015.

[38] N. Karunanayake, “OMR Sheet Evaluation by Web Camera Using Template Matching Approach,” *International Journal for Research in Emerging Science and Technology,* vol. 2, no. 8, pp. 40-44, Aug 2015.

[39] H. Tjahyadi, Y. Budijono, S. Lukas and D. Krisnadi, “Android Based Automated Scoring of Multiple-Choice Test,” *International Journal of Machine Learning and Computing,* vol. 7, no. 5, pp. 110-113, Oct 2017.

[40] P. Sanguansat , “Robust and Low-Cost Optical Mark Recognition for Automated Data Entry,” Hua Hin, 2015.

[41] M. Čupić, “A case study: using multiple-choice tests at university level courses – preparation and supporting infrastructure,” *Int. J. Intelligent Defence Support Systems,* vol. 3, no. 2, pp. 90-100, 2010.

[42] M. Čupić , B. Karla , T. Hrkać and Z. Kalafatić , “Supporting automated grading of formative multiple choice exams by introducing student identifier matrices,” *MIPRO, 2011 Proceedings of the 34th International Convention,* 993-998.

[43] I. Bonačić , T. Herman , T. Krznar, E. Mangić and G. Molnar , "Optical character recognition of seven–segment display digits using neural networks," *Proceedings of MIPRO 2009 – 32nd International Convention on Information and Communication, Technology, Electronics and Microelectronics,* pp. 323-328, 2012.

[44] M. Čupić, K. Brkić , T. Hrkać , Z. Mihajlović and Z. Kalafatic, "Automatic recognition of handwritten corrections for multiple-choice exam answer sheets," *37th International Convention on Information and Communication Technology, Electronics and Microelectronics,* pp. 1136-1141, 2014.

[45] V. Chouvatut and S. Prathan , "The Flexible and Adaptive X-Mark Detection for the Simple Answer Sheets," Khon Kaen, 2014.

[46] D. Chai, "Automated marking of printed multiple choice answer sheets," *IEEE International Conference on Teaching, Assessment, and Learning for Engineering, IEEE,* pp. 145-149, 2016.

[47] G. Bayar, "The Use of Hough Transform to Develop an Intelligent Grading System for the Multiple Choice Exam Papers," *Karaelmas Science and Engineering Journal,* vol. 6, no. 1, pp. 100-104, 2016.

[48] G. Bayar, "Development of an Automatic Grading System Based on Energy Circular Hough Transform and Causal Median Filter," *International Journal of Intelligent Systems and Applications in Engineering,* vol. 5, no. 3, pp. 81-88, 30 Sep 2017.

[49] A. M. Abdu, and M. M. Mokji, "A Novel Approach to a Dynamic Template Generation Algorithm for Multiple-Choice Forms," in *2012 IEEE International Conference on Control System, Computing and Engineering*, Penang, 2012.

[50] M. Afifi and K. Hussain, "The Achievement of Higher Flexibility in Multiple Choice-based Tests Using Image Classification Techniques," *IEEE Learing Technologies,* pp. 1-8, 2017.

[51] W. sheng , S. Wang and H.-l. Xiao , Support Vector Machines for User-Defined Sheets Recognition in Complex Environment, vol. 6839, Berlin Heidelberg: Springer, 2012, pp. 197-202.

[52] A. Patil, M. Naik and P. Chare , “FPGA Implementation for OMR Answer Sheet Scanning Using State Machine and IR Sensors,” *International Journal Of Electrical, Electronics And Data Communication,* vol. 4, no. 11, pp. 15-20, Nov 2016.

[53] F. Zampirolli, V. Batista and J. Gonzalez, “An Automatic Generator and Corrector of Multiple Choice Tests with Random Answer Keys,” in *IEEE Frontiers in Education Conference*, Erie, 2016.

[54] B. Haskins, “Contrasting Classifiers for Software-based OMR Responses,” in *Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference*, Port Elizabeth, 2015.

[55] Gravic, Inc, “Remark Office Optical Mark Rcognition,” Gravic, Malvern, 2015.

[56] Remark, “Remark Office OMR,” Gravic, Inc, 2014. [Online]. Available: http://remarksoftware.com/products/office/. [Accessed 17 Spt 2016].

[57] Gravic, Inc , “Gravic Remark Office OMR: Version 6.0 for Windows Whitepaper,” Gravic, Malvern, 2006.

[58] R. Szeliski, Computer Vision:, Washington: Springer, 2010, pp. 251-276.

[59] “Aesthetically Evolved Images,” ashleymills, [Online]. Available: https://www .ashleymills.com/book/export/html/128.

[60] J. Sachs, “Digital Image Basics,” *Digital Light & Color,* pp. 1-14, 2000.

[61] “Image processing toolbox,” ROHAN Academic Computing, [Online]. Available: http://www-rohan.sdsu.edu/doc/matlab/toolbox/images/color5.html.

[62] P. Li , H. Shen and L. Zhang, “A Method of Image Resolution Enhancement Based on The Matching Technique,” in *Mapping and Remote Sensing*, Wuhan , 2011.

[63] R. Szeliski, “Image Alignment and Stitching,” *Microsoft Research,* pp. 1-44, 27 September 2006.

[64] H. Rhody, “Geometric Image Transformations,” *Chester F. Carlson Center for Imaging Science,* pp. 1-38, 8 Sep 2005.

[65] Puneet and N. Garg, “Binarization Techniques Used for Grey Scale Images,” *International Journal of Computer Applications,* vol. 71, no. 1, pp. 8-11, June 2013.

[66] E. Balamurugan, K. Sangeetha and P. Sengottuvelan, “Document Image Analysis- A Review,” *International journal of Computer application,* vol. 1, no. 1, pp. 20-26, 2011.

[67] Matlab, “Image Processing ToolBox,” Matlab, 2016. [Online]. Available: https://au.mathworks.com/products/image/?requestedDomain=au.mathworks.com&requestedDomain=au.mathworks.com. [Accessed 12 Oct 2016].

[68] R. Singh, S. Roy, I. Singh, T. Sinam and K. Singh, “A New Local Adaptive Thresholding Technique in Binarization,” *IJCSI International Journal of Computer Science Issues,* vol. 8, no. 6, pp. 271-277, Nov 2011.

[69] W. Peerawit and A. Kawtrakul , “Marginal Noise Removal from Document Images Using Edge Density,” in *IEEE Conference on Cybernetics and Intelligent Systems, 2004*, Singapore, 2004.

[70] A. Hambal, Z. Pei and F. Ishabailu, “Image Noise Reduction and Filtering Techniques,” *International Journal of Science and Research,* vol. 6, no. 3, pp. 2333-2038, 2015.

[71] A. Karne and S. Navalgund, “Implementation of an Image Thinning Algorithm Using Verilog and MATLAB,” *International Journal of Current Engineering and Technology,* pp. 333-337, 2013.

[72] P. Bansal and B. Kaur, “A Review on Thinning in Digital Image Processing,” *International Journal of Science and Research,* vol. 5, no. 4, pp. 228-231, 2015.

[73] “zipgrade Papers,” zipgrade, 2017. [Online]. Available: www.zipgrade.com/forms. [Accessed 2018].

[74] . T. Hoffman , “The Best All-in-One Printers of 2016,” PCMag, 22 Sep 2016. [Online]. Available: au.pcmag.com/ printer-reviews /12556 /guide /the-best-all-in-one-printers-of-2016. [Accessed 12 Oct 2016].

[75] RICOH, “Aficio MP c5501A,” 2015. [Online]. Available: www.ricoh.com.au/aficio-MP-C5501a-brochure. [Accessed 12 Oct 2016].

[76] Matlab, “Matlab Is Designed for Engineers and Scientists,” Matlab, 2016. [Online]. Available: https://au.mathworks.com/products/matlab/features.html. [Accessed 12 Oct 2016].

[77] Scantron Technology, “Scantron OpScan 8,” [Online]. Available: http://www.scantron.com/~/media/Scantron/Files/products/scanners/opscan/OpScan%208%20Datasheet.ashx. [Accessed Dec 2017].

[78] J. A. Fisteus, A. Pardo and N. F. Garcı´a, “Grading Multiple Choice Exams with Low-Cost and Portable,” *Springer Science & Business Media,* pp. 560-571, 23 Sep 2012.

[79] Scantron Technology, “Scantron OpScan 8,” Scantron, 2014. [Online]. Available: http://www.scantron.com/~/media/Scantron/Files/products/scanners/opscan/OpScan%208%20Datasheet.ashx. [Accessed Dec 2017].