INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]

FORM REGISTRATION : A COMPUTER VISION APPROACH

by

Ramin Safari-Foroushani

**A Dissertation
Submitted to the Faculty of Graduate Studies and Research
Through the Department of Electrical Engineering
in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy
at the University of Windsor**

**Windsor, Ontario, Canada
1997**

**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

**395 Wellington Street
Ottawa ON K1A 0N4
Canada**

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

**395, rue Wellington
Ottawa ON K1A 0N4
Canada**

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-52413-2

Canada

© **1997 Ramin Safari-Foroushani**

ABSTRACT

FORM REGISTRATION : A COMPUTER VISION APPROACH

by

Ramin Safari-Foroushani

An important problem in office automation is the machine extraction and recognition of filled-in information in form documents. This thesis addresses the extraction of the user entered information in forms, when an original blank form is available (herein after referred to as the *master*) as a scanned document.

Starting with a filled in form (herein after referred to as the *input form*), this thesis addresses the problem of converting the coordinate representation of the input form so that it matches the master. In addition to the data that is filled-in, the input form differs from the master in that it could be a translated, non-uniformly scaled, rotated, or sheared version of the master.

Under normal conditions if the master and the input forms were scanned under similar circumstances, then the input form and the master would be nearly identical except for the filled-in items. In this thesis, no constraints have been imposed on the magnitude of the distortion with regard to translation, scaling, rotation and shear of the input form with regard to the master. However, the proposed approach requires that a master form be available.

The principal problem is the determination of the correspondence between the selected features in both images. The features selected are points on the Convex Hull and straight lines inside the form. It is assumed that the geometric distortion can be modeled as an affine transformation and that point and line features are adequate for establishing correspondence.

The correspondence is solved by applying the Best Find strategy on a set of five points geometric invariants and the Geometric Hashing matching strategy to the four different affine invariants proposed by this thesis. An integration of voting results based on a statistical point of view is performed. Using this correspondence result, the transformation parameters are found by a Least Square solution. The transformation parameters are verified and updated. The input form is mapped to its corresponding template using the transformation parameters obtained.

The performance and robustness of the proposed method is then evaluated using sensitivity analysis of the proposed method and OCR (Optical Character Recognition) analysis of the registered forms.

To my wife Shahrbanoo Hatami for her support.

To my lovely children Nagine and Armin.

To my parents Rajab-Ali Safari and Mehrangiz Alamdari.

ACKNOWLEDGMENTS

I am sincerely grateful to my advisors Dr. M. Ahmadi and Dr. M. Shridhar for their help and support. I would like to thank Dr. N. Narasimhamurti for his impressive mathematical background and new ideas in my research without whom I wouldn't be able to carry on my program.

TABLE OF CONTENTS

ABSTRACT	iv
DEDICATION.....	vi
ACKNOWLEDGMENTS.....	vii
TABLE OF CONTENTS.....	viii
LIST OF TABLES.....	x
LIST OF ILLUSTRATIONS	xi
I. INTRODUCTION	1
1.1 Document Analysis and Understanding	2
1.2 A Complete Form Processing System	5
1.3 Scanner Misregistration.....	8
1.4 Form Processing	9
1.5 Goal Of The Thesis	10
1.6 Thesis Outline	11
II. GEOMETRIC MODELING OF FORMS	12
2.1 Introduction.....	12
2.2 Geometric Transformation.....	13
2.2.1 Projective Transformation	14
2.2.2 Affine Transformation.....	15
2.2.3 Distortions in Input Forms	16
2.3 Geometric Invariants	17
2.3.1 Projective Invariants.....	17
2.3.2 Affine Invariants	19
2.4 Matching strategies	23
2.4.1 Hough Transformation (Pose Clustering).	23
2.4.2 Alignment	24
2.4.3 Geometric Hashing	24
2.4.4 Complexity analysis	26
2.5 Summary and Conclusions.....	26

III. FORM REGISTRATION	28
3.1 Introduction	28
3.2 Feature Extraction	29
3.2.1 Key Point Extraction	30
3.2.2 Line Extraction	32
3.3 Correspondence Determination	37
3.3.1 Projective Model Approach	38
3.3.2 Affine Model Approach	39
3.4 Distortion Evaluation	41
3.5 Registration and Experimental Results	45
3.6 Summary and Concluding Remarks	55
 IV. PERFORMANCE EVALUATION	 57
4.1 Sensitivity Analysis	58
4.1.1 Sensitivity Analysis of the Feature Extraction	59
4.1.2 Sensitivity Analysis of the Affine Invariants	61
4.1.3 Robustness	63
4.1.4 Sensitivity Analysis Remarks	70
4.2 OCR Evaluation	71
4.2.1 OCR Evaluation of the Registered Data	71
4.2.2 Overall Performance Evaluation of the Algorithm	73
4.2.3 Conclusions	79
4.3 Discussions	80
 V. SUMMARY AND CONCLUSIONS	 81
5.1 Summary	81
5.2 Conclusions	82
5.3 Future Research	83
 Bibliography	 85
 Appendix	 92
 Publications	 104
 Vita Auctoris	 105

LIST OF TABLES

Table 4.1	The average and the worst case of the error obtained in transformation parameters.....	69
Table 4.2	The average and the worst case of the distance error obtained in registering a 2000 by 2000 square	70
Table 4.3	Average and variance of the reduction in recognition rate after application of the proposed algorithm compared to the non-distorted templates..	73

LIST OF ILLUSTRATIONS

Figure 1.1	A complete form processing system	5
Figure 1.2	From left to right :Aligned form, sheared form, rotated form, translated form, non-uniformly scaled form.....	8
Figure 2.1	b,c,d,e and f images are different projective transformations of image a	15
Figure 3.1	An example of a Convex subset of a set of pixels (gray pixels form a convex).....	29
Figure 3.2	A sample of a template with some arrows superimposed on the image to show the possible candidates for its Convex Hull.....	31
Figure 3.3	A sample of a template form	34
Figure 3.4	The lines extracted from the template in figure 3.3.....	35
Figure 3.5	The image showing the accuracy of the line detection algorithm	36
Figure 3.6	The rotated input filled in form with 75 % scaling involved	46
Figure 3.7	The registered input filled-in form with scaling and rotation corrected.....	47
Figure 3.8	The Morphological closing result of figure 3.7	48
Figure 3.9-a	First sample of template form	49
Figure 3.9-b	First sample of input form	49
Figure 3.9-c	First sample of registered form.....	50
Figure 3.9-d	Extracted information from fig. 3.9-b	50

Figure 3.10-a Second sample of template form	51
Figure 3.10-b Second sample of Input form.....	52
Figure 3.10-c Second sample of registered form.....	53
Figure 3.10-d Extracted information from 3.10-b.	54
Figure 4.1 A portion of the input form	59
Figure 4.2 100% noise added to Figure 4.1	59
Figure 4.3 Average error in keypoints extracted due to noise added to the form.....	60
Figure 4.4 First form used for the robustness test	65
Figure 4.5 Second form used for the robustness test	66
Figure 4.6 Third form used for the robustness test	67
Figure 4.7 Fourth form used for the robustness test.....	68
Figure 4-8 Average recognition rate of the non-registered forms in the presence of shearing	74
Figure 4-9 Average recognition rate of the non-registered forms as a function of scaling.....	75
Figure 4-10 Average recognition rate of the non-registered forms in the presence of rotation.....	75
Figure 4-11 Average recognition rate of the registered forms as a function of shearing	76
Figure 4-12 Average recognition rate of the registered forms as a function of scaling.....	76

Figure 4-13	Average recognition rate of the registered forms as a function of rotation	77
Figure 4-14	Average improvement in recognition rate of the registered forms in percentage considering the effect of shearing.....	78
Figure 4-15	Average improvement in recognition rate of the registered forms in percentage considering the effect of scaling	78
Figure 4-16	Average improvement in recognition rate of the registered forms in percentage considering the effect of rotation.....	79
Figure A-1	First sample of the blank template	92
Figure A-2	First sample of the filled in form with geometric distortion	93
Figure A-3	First sample of the registered image to its template	94
Figure A-4	Extracted information from figure A-3.....	95
Figure A-5	Second sample of the blank template	96
Figure A-6	Second sample of the filled in form with geometric distortion	97
Figure A-7	Second sample of the registered image to its template	98
Figure A-8	Extracted information from figure A-7.....	99
Figure A-9	Third sample of the blank template	100
Figure A-10	Third sample of the filled in form with geometric distortion	101
Figure A-11	Third sample of the registered image to its template	102
Figure A-12	Extracted information from figure A-11.....	103

CHAPTER I

Introduction

In recent years document image analysis and understanding has become an important research area. Automated processing of large volumes of documents such as bank checks, invoices, technical manuals, forms, etc. can lead to significant improvements in office productivity and related applications. Due to recent advances in this field, manual processing of various types of forms and tables is steadily being replaced by a computer based document processing and recognition system. Applications include income tax returns, revenue documents, census forms, address reading, various application forms, reader cards, cheques and bills, which are widely used and are essential to all business and government institutions. The current state of research in document analysis and understanding is summarized in section 1.1.

A complete form processing system is introduced in section 1.2 and the problem of the existence of the geometric distortion in documents is presented in section 1.3.

Automation is widely demanded where large amounts of filled-in documents must be converted into a computer readable form. This has led many researchers to develop various algorithms for the purpose of form processing and recognition. The current state of the art in the area of form processing is presented in section 1.4. Section 1.5 presents the goal of this thesis. The outline of this thesis is presented in section 1.6.

1.1 Document Analysis and Understanding.

Recently document image analysis/understanding systems have been studied for the purpose of office automation. In these systems, a document image is analyzed and then divided into different categories of information which can be reconstructed as a multimedia document standardized by CCITT and ISO [50].

Several techniques for analysis and recognition of a variety of documents with different layout and languages have been proposed by researchers[32 - 43]. Wahl, Wong, and Casey [32] presented a smearing algorithm which is widely used in documents such as journals and printed materials. Toyoda, Noguchi, and Nishimura [33] extracted Japanese newspaper articles using domain-specific knowledge. Okamoto and Miyazawa [34] analyzed documents containing mathematical expressions. Baird and Thompson [35] studied the extraction of chess games from the chess informant. Esposito, Malerba, and Semeraro [36] classified document types (patents, scientific papers, etc.). Higashino, Fujisawa, Nakano, and Ejiri [37] proposed a flexible format understanding method using FDL, i.e., a format definition language for classification of different blocks in printed materials. Tsuji [38] represented a document structure with tree. Nakano, Fujisawa, and Kunisaki [39] have developed a business form understanding system incorporating character recognition. Inagaki, Kato, Hiroshima, and Sakai [40] have developed a specific-purpose machine for Japanese document understanding. Masuda, Hagita, and Akiyama [41] presented a prototype of a Japanese text reader. Baird [42] proposed a versatile text reader whose layout analysis component was designed to be language-independent. Srihari, and Zack [43] surveyed the current status of document analysis in depth.

The first step in any document analysis system is to correct the geometric distortions that occur during the representation of the image to the system. These geometric distortions consist of the primitive distortions such as rotation (skew), (non-uniform) scaling, shearing, translation and reflection. In the literature several attempts to correct the skew of the input documents [10-14,51-54] have been reported.

A commonly used method for skew detection employs the projection profile to iteratively compute the skew angle [10]. The projection profile is a histogram of the number of black pixel values accumulated along parallel lines throughout the document. The projection profile is computed at a number of angles close to the expected orientation, and for each angle, a measure is made for the total variation in the bin heights along the profile. The maximum variation corresponds to the best alignment with the text lines, and this determines the skew angle. This method performs very well in cases where the information inside the document is printed text with minimal graphics and variation in font size.

Baird [11] proposed modifications to the projection profile method for very fast and accurate iterative convergence on the skew angle. Akiyama and Hagita [12] describe an approach that is very fast but less accurate. A document is divided into equal width vertical strips; horizontal projection profile is calculated for each vertical strip, and skew is determined by measuring the average shift in zero crossing between strips. Pavlidis and Zhou [51] proposed a method where the alignment is measured by vertical projection profiles determined for horizontal strips of the document. A restriction common to these

methods is that they are limited to documents that have relatively small skew distortion (less than 10 degrees in magnitude).

Srihari and Govindaraju [52] and Hinds, Fisher and D'Amato [14] proposed similar skew detection methods based on the Hough transform. In this method, each black pixel is mapped to the Hough space and the dominant peaks in the Hough space determine the skew angle of the document. The limitation of this approach is that if the text is sparse, it may be difficult to correctly choose a peak in Hough space. The specified skew angle is limited to 15 degrees in magnitude and it is very sensitive to the existence of landmarks and graphics in the document.

There are other techniques reported in the literature to detect the skew angle such as the nearest-neighbor clustering approach proposed by Hashizume, Yeh, and Rosenfeld [53]. In this method, the 1-nearest-neighbors of all connected components are found, the direction vectors for all nearest-neighbor pairs are accumulated in a histogram, and the histogram pick is found to determine the dominant skew angle. This method has the advantage over the previous methods in that it is not limited to any range of skew. However, since only 1-nearest-neighbor connection is made for each component, connections with noise, subparts of characters (dot on "i"), and between-line connections can reduce the accuracy of this method. This approach has been improved by O'Gorman [54] yielding more accurate measure of the skew and also introducing other layout structures of the document such as within-line and between-line spacing and locating text blocks and text lines inside the document.

1.2 A Complete Form Processing System.

A complete form processing system has several key components which are presented in figure 1.1.

Figure 1.1 A complete form processing system.

The form processing task should normally start with a well designed form that is both easy to use and easy to read. In environments where the processing is done by an operator, the design principles which make a form attractive and easy to complete typically give rise to a form which is reasonably easy to process. In an automated environment, however, the design is normally motivated by the need to automate the processing of the form. For example, if the form is printed in red ink, a filter can be used during the scanning process to eliminate the form components and simplify the information extraction process. Similarly, the form can incorporate a code for easy form identification, and landmark features strategically placed to aid skew detection and alignment. If the automatic form processing system uses a geometric modeling of the form structure, the form design issue can be guided by using some geometric properties of the appropriate shapes or locations of the geometric features on the form. The problem with the form design is that forms which are ideal for automated processing may be prohibitively expensive to produce or simply unattractive to the user. Our goal is to provide a form processing environment which allows greater flexibility in the types and numbers of forms which are to be processed. The proposed technique assumes the availability of a sample of the pre-designed form and therefore the design of forms are not addressed in this thesis, however, there is a potential for improving the overall approach if the forms are pre-designed according to specific needs.

The proposed form processing environment should be able to process input filled-in forms with the same layout given. The task is to use a blank sample of the form to be processed as a template to register any input filled-in form. This process is called form

registration and the result should be the removal of any typical geometric distortion that may happen during the scanning process. After correctly registering the input filled-in form to its corresponding template, any filled-in information is now mapped to its corresponding blank data field in template form. These fields are then recovered and passed to a storage or Optical Character Recognition (OCR) system. Based on prior knowledge about the type of information filled-in, one has to use either a handwritten recognition system, a printed text recognition system or both.

In this thesis, a geometric approach using selected point and line features is adopted to develop an analytic model for the input forms. The proposed method does not require any restrictions on the form to be processed; however a blank sample of the form is assumed to be available as a reference. This blank sample is processed and a geometric model is derived and stored as a reference. An affine transformation is used to characterize the distortion that occurs as a filled-in form is scanned and digitized. This requires that:

- (i) relevant features inside the form are easily extracted
- (ii) geometric affine invariants are derived using these features.

Then a geometric hashing voting scheme is used to find the correspondence between the relevant features in the filled-in form and the blank sample form(template). The transformation parameters required to map the filled-in form to the template are found using these correspondences.

1.3 Scanner Distortion.

Scanners are normally used to digitize and store form images on a computer. Several distortions are introduced during the scanning process. While some of these distortions can be reduced by careful hardware design or corrected by pre-processing techniques in early stages of form processing systems, certain types of distortions such as scaling, multi-dimensional shearing and quantization noise, etc. are not completely eliminated.

Distortions that are not easy to correct include the differences between the geometrical features of a digitized filled-in form and the corresponding scanned blank sample of the form. These can be the differences in offset (horizontal and/or vertical translation), rotation (horizontal skew angle of the document), scaling (either in x and/or y direction) and shearing (which can be considered as rotation followed by a non-uniform scaling) or any combination of these distortions. Figure 1.2 displays some typical geometric distortions which may be present in a scanned document. In this figure the outer boundary represents the viewing area of the scanner.

Figure 1.2 : From left to right :Aligned form, sheared form, rotated form, translated form, non-uniformly scaled form

The form registration approach, proposed in this thesis, attempts to correct these geometric distortions from the scanned input filled-in form and to map the form to its corresponding undistorted blank template for extraction of filled-in information .

1.4 Form Processing.

The development of an automatic form reading machine is critical to effective office automation. There are essentially two critical steps that need to be addressed in some detail. The first step involves the registration of the filled-in input form to an aligned form which is free of geometric distortions. The second step deals with the extraction of the filled-in information from the registered form for recognition and storage. This thesis concentrates on the first task that addresses form registration.

Recently, a general method for extracting the filled-in information for various kinds of forms has been proposed by Lei Xu [44,48]. This method is based on a Generalized Randomized Hough Transform (GRHT) [49] to find a rotation parameter from the scanned filled-in form. Using this parameter, the filled-in form and the blank sample of the form are matched and uncommon parts which represent the filled-in information are extracted. However, this method is computationally intensive when the filled-in form is skewed at angles exceeding 10 degrees in magnitude. Casey and Fergusen [45] described a form processing system which allows model definition, data extraction and conversion to a symbolic form for efficient storage in a database. The areas in which data are located is defined by the form. In this technique, translation and rotation are adjusted. The system recognizes the existence of characters intersecting with form line components and

proposes a solution in which the pixels in the intersection are not deleted. The shortcoming of this technique is also the limited range allowed for the skew to happen. Taylor, Fritzson, and Pastor [46] described a system for modeling, identifying and processing IRS forms. The problem of non-form/form interaction is not addressed in this paper. Wang and Srihari [47] described an approach which does not require any a priori knowledge of the form structure. They decompose the form into boxes, line segments, logos, and filled-in data. Characters broken by line removal are joined using a character patching method. The input filled-in form assumed to have no geometric distortion such as skew, etc. This approach, though less restrictive than many existing techniques, would definitely benefit greatly from additional knowledge about the original form.

1.5 Goals of the Thesis.

Literature survey in the area of document image processing has shown several effective algorithms and techniques developed for automated processing of printed documents such as newspapers, magazines, manuals etc. In the area of form processing, however, very few systems capable of automatic processing and recognition have been developed. As the demand for automated form processing increases, significant new research directions have been proposed by the researchers to improve the current techniques. Until recently, form processing systems assumed the input form to have only geometric distortions such as rotation (less than 10 degrees in magnitude) and translation. However, experimental studies have shown that these are not the only possible distortions and there might be some other types of distortions such as non-uniform/uniform scaling,

shearing, and rotation more than 10 degrees in magnitude (specially for small size forms). This thesis, therefore, will address the form registration task of a complete form processing system. The methods proposed are capable of correcting many types of geometric distortions (with relatively few restrictions) in input forms. The proposed methods use a computer vision approach for form modeling and registration, using a blank sample of the form as template (model). Experimental studies with a variety of form documents have shown the method to be reliable and promising.

1.6 Thesis Outline.

The geometric modeling of an object in computer vision will be presented in chapter II. This chapter covers the invariant pattern matching issue in computer vision followed by different strategies available for matching between geometrically related objects.

A new form registration method is presented in chapter III. The performance of the proposed method is evaluated and presented in chapter IV, covering the step by step sensitivity analysis of the proposed method followed by an OCR evaluation of the registered data and overall performance analysis of the proposed method based on an OCR point of view. Chapter V presents some relevant conclusions and summarizes contributions of this thesis.

CHAPTER II

Geometric Modeling of Forms : A Computer Vision Approach

2.1 Introduction.

The form registration process described here is based on a model based recognition technique that is often adopted by computer vision researchers. Object recognition is a major task in computer vision. In most practical applications we may assume that the objects to be recognized are known in advance. This model-based approach proved to be promising, and the best currently known object recognition systems are model-based. (for comprehensive survey see [22,23]). In this method, one derives a geometrical model of the input image. This model extracts an invariant signature from relevant features of the image.

In order to discuss object recognition algorithms, one has to start with a mathematical description of an object. The object can be described using many attributes, such as its convex hull, surface curvatures, edges, point and/or line features, etc.

In one model-based object recognition algorithm, described as *Pose Clustering* or *Alignment* [17-19] features extracted from a test object are matched against features of the model. Alternatively one can use a *geometric hashing* technique [20-21] in which each subset of relevant features is represented by an invariant signature, and the task is then to find the correspondence between feature subsets in the model and the test object.

The form processing problem can be tackled in a similar manner. Assuming that a blank sample of a given form is available, one can construct a geometrical model of the

form using geometric features inherent to the model. Features extracted from a filled-in form are then matched against those of the model to facilitate the extraction of user entered information. In section 2.2, two different transformations that can be used as a model for form registration applications, are presented. Geometric invariants are discussed in section 2.3. Different matching strategies reported in literature will be critically compared in section 2.4. Section 2.5 presents relevant conclusions and observations.

2.2 Geometric Transformation.

Transformation occurs in a variety of mathematical formulations of a problem and often implies a change of scale or a redefinition of one or more variables in an equation or expression to facilitate further analysis and interpretation. Alternatively, one may use “transformation” to define a functional relationship between geometric objects. It is customary to call a functional relation a **transformation** or **mapping**. The geometric objects that play the role of independent variables are called **originals** or **models**, and those constituting the dependent variables are called the **correspondents** or **images**.

Here, transformation is used to map one set of features to another set of similar or different features. For example a set of points can be mapped or transformed to a set of lines. We shall be concerned mainly with **point transformations**, which involve transforming point features in model to point features in images.

Among the many transformations that could be defined, the **projective transformation** and the **affine transformation** which are of interest in form registration applications will be discussed. In the context of form processing the mapping will address

the extraction of the transformation parameters relating the blank form and the filled-in test form.

2.2.1 Projective Transformation.

A projective transformation of a plane onto itself is a transformation which

- a) sends a plane onto itself in a one to one correspondence.
- b) preserves Collinearity.

Every projective transformation that maps a point (x, y) to another point (x', y') can be presented as follows :

$$x' = \frac{a_1x + a_2y + a_3}{c_1x + c_2y + c_3}, \quad y' = \frac{b_1x + b_2y + b_3}{c_1x + c_2y + c_3} \quad \text{where } \Delta = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \neq 0 \quad (2.1)$$

Conversely, such equations always represent a projective transformation.

It is well known that there is a unique projective transformation that sends four given points into another four given points, respectively. if no three points of either set are collinear [9]. One can conclude that the transformation parameters can be obtained from at least four point correspondences between two projective transformed images. Two images are said to be **projectively equivalent** if either image can be mapped onto the other by a projective transformation. The following figure 2.1 shows a number of projectively equivalent images that are related to each other by some specific projective transformation. Each projective transformation can be decomposed into rotation, translation, scaling in x and/or y direction, shearing, and the perspective distortion.

Figure 2.1 : b, c, d, e, and f images represent different projective transformations of image a.

2.2.2 Affine Transformation.

An affine transformation is a subclass of the projective transformation which maps the plane on to itself and further the inverse mapping is also affine. An affine transformation of a point (x, y) into (x', y') can be represented as follows

$$\begin{cases} x' = ax + by + m \\ y' = cx + dy + n \end{cases}, \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0 \quad (2.2)$$

It is clear that affine transformation is a linear transformation that maps the plane into itself. Two images are said to be **affinely equivalent** if and only if there is an affine transformation that maps one image into the other.

There is a theorem in affine geometry which states [9]:

There is a unique affine transformation which maps any three non-collinear points into any other three non-collinear points respectively.

One can conclude that the affine transformation parameters can be obtained starting from at least three point correspondences between two affine transformed images.

The affine transformation is a subclass of projective transformation without the possibility of perspective distortion. Each affine transformation can be decomposed into the primitive distortions such as rotation, scaling in x and or y direction, shearing, and translation.

2.2.3 Distortions In Input Forms.

In this thesis, we are using a blank sample of the form to be processed as a **model** or **template**. In addition to the data that is filled-in, the input form differs from the template as it may have become a translated, (non-uniformly) scaled, rotated or sheared version of the template. Each one of these distortions is a primitive geometric transformation. Every affine transformation can be decomposed into the above primitive distortions. Assuming that global distortion can happen under the scanner, the affine transformation would be the best choice for the distortion model.

One can also use the projective transformation for the distortion model. The reason is that the affine transformation is a subclass of the projective transformation which does not cover the perspective distortion. The projective transformation covers all the primitive distortions mentioned above plus the perspective distortion. The experimental results show that the affine transformation and projective transformation are the best choices assuming that a global distortion occurring under the scanner. In this thesis, we have used both models for the geometric distortion in input filled-in forms.

2.3 Geometric Invariants.

In computer vision, there are a variety of methods used to derive the transformation parameters that maps a given image into its corresponding model. In this thesis, we use an invariant based matching approach to extract the parameters [1..4]. The first step would be to choose a transformation that best fits the distortion model. Next step would be to derive unique features from a given image and a corresponding model. In order to find the correspondences between features in two images, one needs to define invariants based on the transformation chosen. This section describes how one can utilize an invariant signature on a given set of features. There has been a variety of invariants proposed by the researchers in literature [7,8,16]. The proposed invariant signature should be invariant under the selected transformation. Affine transformations, for example, are concerned with affine invariants, while projective transformations deal with projective invariants, etc. In the two following sections, a geometric invariant (Cross Ratio) based on a set of five points, no three of them collinear, and four different affine invariants based on keypoint and line features is presented.

2.3.1 Projective Invariants.

In this section, we provide some basic results on the effect of rotation, translation, scaling of the axes, and shearing on a set of points in a plane. Rotation, translation, scaling of the axes, and shearing are special cases of general projective transformations. A fundamental result in projective geometry states that :

If four points A, B, C, and D, are such that no two of them are collinear with a fifth point P, then the cross ratio of the four points with respect to P, defined as

$$\lambda = \frac{[AD][BC]}{[AC][BD]}, \text{ where } z \perp \text{ plane, } [AD] = z \bullet (A - P) \times (D - P) \text{ etc.} \quad (2.3)$$

is invariant under any projective transformation [6].

In the above equation each point is represented by its homogenous coordinate representation $(x, y, 1)^T$, $z = (0, 0, 1)^T$, $\bullet$ represents an inner product operation and $\times$ represents a cross product operation between the two vectors.

The cross ratio λ defined in (2.3) depends on the order of the points A, B, C, and D. However, the 24 possible permutations of the four points give rise to only 6 cross ratios:

$$\lambda, \frac{1}{\lambda}, 1 - \lambda, \frac{1}{1 - \lambda}, \frac{\lambda}{1 - \lambda}, \frac{1 - \lambda}{\lambda} \quad (2.4)$$

The following parameter which is the smallest algebraic ratio denoted by $\rho(A, B, C, D: P)$ would be permutation insensitive :

$$\rho(A, B, C, D: P) = \text{algebraic minimum} \left(\lambda, \frac{1}{\lambda}, 1 - \lambda, \frac{1}{1 - \lambda}, \frac{\lambda}{1 - \lambda}, \frac{1 - \lambda}{\lambda} \right), \quad (2.5)$$

where $\lambda = \lambda(A, B, C, D: P)$

Given five points **a, b, c, d** and **p** in the template form, and five points **A, B, C, D,** and **P** in the input form, one needs to answer the following question :

*Under what conditions can one conclude that there is a projective transformation that maps **a** with **A**, **b** with **B**, and so on?*

If $\lambda(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}: \mathbf{p}) \neq \lambda(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}: \mathbf{P})$ then the answer is clearly no. However, if the cross ratios are equal, we cannot conclude that such a transformation exists. In fact, if

we fix A, B, C, P , and λ , the locus of D that satisfies equation (2.3) is a straight line passing through P . Thus the equality of cross ratios is necessary but not sufficient. However, the following theorem gives the necessary and sufficient conditions for an affirmative answer.

Theorem: Suppose that a, b, c, d and e are any five points in the template, and A, B, C, D and E are any five points in the document; then there is a projective transformation that maps a with A , b with B , and so on if and only if $\lambda(a, b, c, d: e) = \lambda(A, B, C, D: E)$, $\lambda(b, c, d, e: a) = \lambda(B, C, D, E: A)$, $\lambda(c, d, e, a: b) = \lambda(C, D, E, A: B)$, $\lambda(d, e, a, b: c) = \lambda(D, E, A, B: C)$, and $\lambda(e, a, b, c: d) = \lambda(E, A, B, C: D)$.

The following statement represents the necessary and sufficient condition for affine transformation as a link between projective geometry and algebra.

Suppose that a, b, c, d and e are any five points in the template form and A, B, C, D , and E are any five points in the input form, then there is a projective transformation that maps a with A , b with B , and so on if and only if the cross ratios of all possible circular permutations of this set of five points are equal.

One can determine the correspondence between a and A , b and B , etc. using the invariant signature (λ or ρ) provided for each set of five points in both images.

2.3.2 Affine Invariants.

In this section, we provide some basic results on the effect of applying an affine transformation on a set of points or lines in a plane. Each point in homogenous coordinate

can be represented as $(x,y,1)^T$ and each line $\{ax+by+c=0\}$ as $(a,b,c)^T$. Each affine transformation in a plane can be represented in matrix form as :

$$X' = TX \quad \text{where} \quad T = \begin{bmatrix} \text{aff1} & \text{aff2} & \text{aff5} \\ \text{aff3} & \text{aff4} & \text{aff6} \\ 0 & 0 & 1 \end{bmatrix}, \quad X = [x \ y \ 1]^T, \quad X' = [x' \ y' \ 1]^T$$

$$L = T^T L' \quad \text{where} \quad L = [a \ b \ c]^T \quad \text{and} \quad L' = [a' \ b' \ c']^T, \quad (2.6)$$

L and X are affine transformed to L' and X'

The key problem is to find the transformation parameters based on two sets of features in a plane. This problem is solved by finding the corresponding points using the invariant signatures. We present four different affine invariants based on key-point and line features.

There is a fundamental theorem in affine geometry with regard to uniqueness which states that [9]:

There is a unique affine transformation that sends three non-collinear points X_1, X_2, X_3 into three non-collinear points X'_1, X'_2, X'_3 respectively.

From this theorem one can conclude that at least four pairs of points are needed in order to determine the pair-wise correspondence between two sets of points. Having four points X_1, X_2, X_3, X_4 , no three of them collinear, one can define an affine invariant μ as the result of an affine transformation of X_1 using the transformation parameters that maps X_2, X_3, X_4 into $(0,0), (1,0)$, and $(0,1)$ respectively.

$$\mu(X_1; X_2, X_3, X_4) = (\alpha, \beta)^T \quad \text{where} \quad \alpha = \frac{\Phi(X_1, X_4, X_3)}{\Phi(X_1, X_2, X_3)} \quad \text{and} \quad \beta = \frac{\Phi(X_1, X_2, X_4)}{\Phi(X_1, X_2, X_3)}$$

$$\Phi(X_1, X_2, X_3) = |Z, X_1, X_2| + |Z, X_2, X_3| + |Z, X_3, X_1| \quad (2.7)$$

$$\text{where } |Z, X_i, X_j| = \det \begin{bmatrix} Z & X_i & X_j \end{bmatrix}, \quad Z = [0 \ 0 \ 1]^T$$

$\Phi(X_i, X_j, X_k)$ is directly related to the area of a triangle with X_i , X_j , and X_k vertices and in affine transformation the ratio of the areas are invariant.

The parameter μ is invariant under any type of affine transformations, so for any T if $X'_i = TX_i$, $i = 1, 2, \dots, 4$ then $\mu(X_1 : X_2, X_3, X_4) = \mu(X'_1 : X'_2, X'_3, X'_4)$.

It should be noted that μ is permutation sensitive to X_2, X_3, X_4 for a fixed X_1 , therefore every permutation of these three points should be considered to define a new affine coordinate.

There is another theorem in affine geometry that states [9]:

An affine transformation multiplies the area of all triangles by the value of its determinant.

From the above theorem, one can conclude that the ratio of the areas remain unchanged under affine transformation. Consider parallel lines $L_i : \{a_i x + b_i y + c_i = 0\}$ represented by $(a_i, b_i, c_i)^T$ and points P_i in a plane represented by $(x_i, y_i, 1)^T$ which are the homogenous coordinate representation. The following two affine invariants can be defined:

Two parallel lines L_i and L_j and one point not located on either of the lines.

$$\delta = \frac{a_i x + b_i y + c_i}{a_j x + b_j y + c_j} \text{ for point } P \text{ and lines } L_i \text{ and } L_j \quad (2.8)$$

Two points P_i and P_j and one line L . (the line connecting P_i and P_j should not be parallel to L)

$$\gamma = \frac{ax_i + by_i + c}{ax_j + by_j + c} \text{ for points } P_i \text{ and } P_j \text{ and line } L \quad (2.9)$$

Definition 2.1. Let A , B and C be three different collinear points. The *division ratio* of the points A , B and C (denoted by (ABC)) is defined as:

$$(ABC) = \frac{\overline{AC}}{\overline{CB}} \quad (2.10)$$

where $\overline{AC}$ and $\overline{CB}$ are defined as the Directed Euclidean Distances of the two points (that is $\overline{AC} = -\overline{CA}$). There is a theorem in affine geometry which states [6] :

The division ratio is affine invariant.

For three parallel lines L_i , L_j and L_k intersecting either x or y axis at points A , B and C , one can define the following affine invariant.

$$\eta = (ABC) = \frac{\overline{AC}}{\overline{CB}} \quad (2.11)$$

The above parameters are invariant under any types of affine transformations. For a given set of features of lines and points the corresponding invariant remains invariant under any type of affine transformation. One can determine the correspondence between key point and line features in two affine transformed images using the invariant signatures (μ , δ , γ , or η) provided for each corresponding feature set selected in both images.

2.4. Matching Strategies.

There are a variety of methods available in literature finding the correspondence between two feature sets belonging to both object and model when the object is a mapping of the model by some specific transformation.

Hough transformation (Pose Clustering) [18], *Alignment* [19], and *geometric hashing* [20] are the most popular and efficient techniques among the others. In this section, we explain briefly the *Hough Transformation* and *Alignment* techniques. Then the *geometric hashing* technique is described in detail since this thesis utilizes this technique. Then a complexity analysis is presented to show the time efficiency of the *geometric hashing* technique compared to the other two.

2.4.1 Hough Transformation.

In the *Hough Transformation (Pose Clustering)* [18] approach, the recognition of an object is achieved by finding a transformation between a model and the object, which maps a large enough number of the model relevant features into object relevant features. The transformation is determined by voting for its parameters.

Let us consider the example of object recognition under the affine transformation (rotation, translation, scaling and shearing) using the Hough paradigm. Assume that both model and object are represented by unique points. Since an affine transformation is uniquely determined by the correspondence triplets of points, each correspondence of a triplet of model points with a triplet of object points provides a candidate affine transformation. An affine transformation can be represented by six independent

parameters. Hence, the Hough transformation technique regards an affine transformation as a point in the 6-D space, each parameter describing a coordinate of this 6-D point.

Although the parameters may achieve a continuum of values, each coordinate axis is quantized into a finite number of values. The final judgment is made based on the maximum vote on the resultant 6-D Hough space.

2.4.2 Alignment.

In the *Alignment* technique, a transformation is hypothesized based on the correspondence of minimal sets of features in both model and object. Then, the candidate correspondence is verified with respect to the other features.

Let us consider the case of affine 2-D matching [19] and assume that both the model and object are represented by unique points. As in the Hough technique, one picks a triplet of points in the object and a triplet point in the model, and computes the affine transformation that maps one triplet into the other. Given a candidate affine transformation, all the other model points are transformed and matched with the object points. If the number of matching unique points is above a certain threshold, the final verification is taken place for the candidate affine transformation parameters.

2.4.3 Geometric Hashing.

The geometric hashing technique for model-based recognition was introduced by Lamdan, Schwartz and Wolfson [20,24,25]. In every model-based recognition system one

has to address two major interrelated problems, namely, object representation and matching.

The objects are represented as a collection of geometric features, such as points and/or lines, and their geometric relations are encoded using the geometric invariants under selected geometric transformation. For example, in the projective transformation, one can use the Cross Ratio, which is a projective invariant in order to geometrically represent each set of five points belonging to the unique points of object or model. For a given model, the first step is to extract the relevant features. Then, based on the transformation and the geometric invariant selected for the problem, one can start building a hash table. Each record in a hash table is a geometric invariant calculated based on a selected feature set. Each entity of the hash table is indexed by its corresponding feature set, so each object can be represented by its own hash table indexed by all possible subsets of its relevant points and/or lines. The hash table preparation stage may be viewed as a learning stage of the algorithm. The most interesting thing about the geometric hashing technique is that the hash table for the model can be built off-line. For each object to be recognized, one has to build a hash table on its relevant features as mentioned before.

The matching stage of the algorithm uses the hash tables prepared in representation stage. In searching through the entities of the object and model hash tables, one can set up a predefined threshold for the relative error in matching a pair of hash table entities in both model and object hash tables. If the relative error is less than a preset threshold, the vote correspond to both feature sets will be incremented in a voting matrix.

The resultant voting matrix has n (number of relevant features in object) rows and m (number of relevant features in model) columns. Each entity (i,j) of this matrix represents how many times the relevant feature i in object considered as a good match to the relevant feature j in model.

Starting from the corresponding features of the highest vote, one can find the transformation parameters required to map the object to the model. The transformation parameters are verified and updated using the next following highest vote.

2.4.4 Complexity Analysis.

There is a comprehensive study by J. Wolfson [21] on the model-based geometric hashing. In this thesis, a brief conclusion of his work on the complexity analysis on geometric hashing, Alignment, and Hough Transformation techniques is presented.

To simplify the analysis, one can assume that the relevant features in object and model are n and m point features. For N number of models and the affine group, the recognition time of the Hough technique is $O(N \times m^3 \times n^3)$, for the Alignment is $O(N \times m^4 \times n^3)$, and for the geometric hashing is $O(n^4)$.

2.5 Summary and Conclusions.

In this chapter, the object recognition task in computer vision has been presented. Two different geometric transformations namely affine and geometric transformation have been discussed. The geometric distortion model has been chosen to be either an affine or the geometric transformation. The Cross Ratio is used as an invariant signature presented

in 2.3.1 for each set of five points assuming a projective transformation as the distortion model, and four different affine invariants proposed in 2.3.2 based on a collection of point and line features assuming an affine model.

We have also concluded that amongst many ways of the methodologies presented such as Hough transformation, Alignment and geometric hashing, the geometric hashing is the least computationally complex algorithm, since in our application the m is equal to the n . We use a geometric hashing strategy to find the correspondence between line and point features in template and input form. The detail of the proposed form registration strategy is presented in the next chapter.

CHAPTER III

Form Registration

3.1 Introduction.

In this chapter a new form registration method is presented. A blank sample of the form (**template**) is chosen as the **model** and a filled-in sample of the form (**input form**) is considered as an **image**. In addition to the data that is filled-in, the input form may differ from the template due to distortion caused by translation, non-uniform scaling, rotation and/or shearing.

To extract the filled-in information from the input form, one has to address the form registration problem. The proposed form registration method can deal with several types of distortions such as rotation, scaling, shearing, non-uniform scaling, and reflection, provided a blank sample of the form is available. These images are scanned as digitized black and white images. The problem of extracting the relevant features in both images is discussed in section 3.2. Section 3.3 addresses the problem of correspondence determination between features in the model and the image. The evaluation of the distortion parameters, required to map the input form to its corresponding template will be discussed in detail in section 3.4. Section 3.5 presents the final registration step and some experimental results. The summary and concluding remarks are given in section 3.6.

3.2 Feature Extraction.

The choice of features used in the form registration algorithm is of critical importance. The selected features in the template and input form should have the following characteristics:

- 1) easily identifiable.
- 2) describe the geometry of the image.

In this thesis, the following features are extracted from the input forms :

- 1) Convex Hull subset of the forms.
- 2) Straight lines inside the forms.

Definition 3.1. *A Convex Hull of a black and white image is a subset of the image which forms a Convex polygon enclosing the whole image.*

Figure 3.1 depicts an example of a set of pixels in a plane and their Convex Hull represented by gray pixels. As can be seen all other pixels are enclosed by the convex polygon formed by the Convex Hull subset of the image.

Figure 3.1. An example of a Convex subset of a set of pixels (gray pixels form a convex).

In form documents, one can always rely on the availability of line features and some key points. Simulation results have shown that if the selected key points are located on the Convex Hull of the image, more accurate results are obtained for the transformation parameters [15]. The affine transformation has the property of mapping a convex set of points to a convex set of points in a plane. Another advantage of using the convex subset of the image is the lack of any restriction on the structure of the forms to be processed. The details of the convex hull algorithm is presented in section 3.2.1.

Line features are extracted using a modified Hough Transform technique [28]. This modification consists of a connected component analysis on the Hough space maxima. The affine transformation has the property of mapping parallel lines into parallel lines in a plane. The details of the line detection algorithm proposed are presented in section 3.2.2.

3.2.1 Key Point Extraction.

Reliable identification of the key points is very critical to the performance of the registration algorithm. In addition, the set of key points must have sufficient asymmetry so that false correspondences are avoided. To increase the accuracy of the resultant transformation parameters, the key points should be away from the center of the form and spread around the form as far as possible. This algorithm uses the points whose convex hull contains the entire image as the set of key points. These points are found by locating the first and last black pixels in each row and column, if they belong to a connected component of at least 16 black points. These points constitute an ordered set of non-

intersecting polygons. The algorithm uses this set of points as the input to the convex hull detection algorithm proposed by Sedgewick [26] to find the convex subset of the form to be processed. Figure 3.2 shows the possible candidates for the key points to be selected as convex hull on a sample form.

• If you reported self-employment income on lines 135 to 143 of your return, complete this schedule. •

You have to file a statement of income and expenses with your return.

Enter the salary or wages you paid in the year to your spouse (if you are in a partnership, enter only your share). _____

Complete the following for the business with the largest total income:

<p>Employer's remittance account Business name _____</p> <p>Location Number Street P.O. Box No. or R.R. No. _____ City _____ Province or territory _____ Postal code _____</p> <p>Briefly describe the major business activity: _____</p> <p>Has the major activity changed since you filed your last return? No ☐ Yes ☐</p>	<p>If the major business activity involves the resale of goods, indicate how you sell these goods: ☐ Retail ☐ Wholesale</p> <p>If your major activity is trucking, are you also an "owner-operator," "leased-operator," or "broker-operator" working for a trucking concern? No ☐ Yes ☐</p> <p>Specify the principal product(s) mined, manufactured, or sold, the type of construction activity or service provided, and the approximate percentage that the revenue from each is of the total revenue.</p> <p>1. _____ % 2. _____ % 3. _____ %</p>
--	---

(Don't forget to complete Schedule 9 on the other side of this page.)

Figure 3.2 : A sample of a template with some arrows superimposed on the image to show possible candidates for its Convex Hull.

This algorithm does not accept noisy black pixels as key points if they do not belong to a connected component of at least 16 pixels. The information written inside the form does not have any effect on the key points extracted as long as they are inside the form, in spite of the possibility of having some written scripts on the boundary of the form

to be processed. In these cases, there might be some key points added or occluded from the set of unique key points to be considered as correct relevant points.

The algorithm has been modified to select at most 8 key points on the convex hull subset of the form which are as far as possible from each other. The advantage of this approach is that one does not require any special marks on the forms. This algorithm has been found to be reliable and less sensitive to noise in the input forms. The detailed sensitivity analysis of this approach can be found in section 4.1.1.

3.2.2 Line Extraction.

There are a variety of methods available in the literature concerning line extraction algorithms from noisy images[27,28,29]. These algorithms map each pixel in the image to a sinusoidal curve in Hough Space ((ρ, θ) space) using the following relationship :

$$\rho = x.\cos(\theta) + y.\sin(\theta) \quad (3.1)$$

where each pixel is represented by its Cartesian coordinate (x,y) .

Each time a sinusoidal curve intersects another at a particular value of ρ and θ , the likelihood increases that a line corresponding to the parameters ρ and θ is present in the original image. An accumulator matrix (consisting of quantized values of ρ and θ) is used to count the number of intersections for various values of ρ and θ [17]. Those cells in the accumulator that have the highest number of votes will indicate possible lines in the original image. This thesis uses the term “possible“ because the high vote may correspond to a text line or graphic. The basic computational steps in the Hough transform are as follows :

for (x)

```

for (y)
  if the pixel is black {
    for (θ) {
      calculate ρ
      increment accumulator array cell (ρ,θ)
    }
  }

```

In this method, there is a trade-off between computational time and the resolution in quantization set up for ρ and θ . Increasing the ρ, θ resolution increases the computational time which in turn increases the accuracy of line detection. In this thesis a resolution of 1° for θ and 8 pixels for ρ is considered when the original image for this operation has been down-sampled by a factor of two. To increase the accuracy of the modified algorithm and to avoid accepting spurious text lines and graphics as valid lines, this algorithm performs a connected component analysis on the components with the highest and next highest votes in Hough space. The segment corresponding to the highest vote in (ρ, θ) space is selected as a valid line if it has a length larger than $2/3 \times \min(\text{width}, \text{height})$ of the image. Width and height are defined as the number of pixels contained in the horizontal and vertical axes of the image respectively. A least square line fitting algorithm applied to a selected connected segment of the image yields the parameters of the line. Each line $\{ a.x+b.y+c=0 \}$ is represented by its homogenous coordinate representation $(a, b, c)^T$.

This algorithm selects the first 8 lines detected by the modified Hough transform approach as line features. Figures 3.3 and 3.4 present an example of the line extraction of the algorithm. The original form image followed by its lines detected and superimposed on a blank canvas is shown in these figures. It should be noted that the spurious peaks in

Figure 3.4 : The lines extracted from the template in figure 3.3.

3.3 Correspondence Determination.

The determination of correspondence between features in the model and the image is crucial to the success of the registration problem. The methods used should be robust enough to cover all problems which may occur during the modeling of the input form such as addition and/or occlusion of relevant features selected in a noisy environment. The algorithm should also avoid false correspondences which often occur during this phase when the relevant features are located symmetrically.

Starting with the features extracted from the template and input form, the next step would be to find the correspondence between the features in both images. This information is then used to find the transformation parameters required that maps the input form to the template. In this thesis two different approaches have been taken to address this problem. In the first approach projective transformation is used to model the distortion. In this approach, the algorithm uses point features as relevant features on the image. These points are located on the convex hull of the image. In this approach, the cross ratio has been used as an invariant signature for each set of five points. The correspondence is determined using the "Best Find" strategy based on a relative error measure between two invariant signatures extracted from both images.

The second approach assumes an affine transformation for the distortion model. In this approach both point and line features have been used as relevant features located on both template and input form images. Four different hash tables are built using the four affine invariants presented in chapter II. The hash tables for the template are built off line and will be used for correspondence determination. The proposed method uses a

geometric hashing strategy to determine the correspondence between feature sets in both template and input form. The voting results obtained from the geometric hashing on four affine signatures are combined based on a statistical model. The final judgment on corresponding features is carried out based on this final result.

3.3.1 Projective Model Approach.

In this approach, the points located on Convex Hull of the template and the input form have been used for matching purposes. The model for the distortion assumed to be a projective transformation model. The invariant signature chosen is the Cross Ratio of each set of five points. Each set of five points from an image gives five cross ratios $(\lambda_i)_{i=1,2,\dots,5}$ as in equation 2.3 computed as the cross ratio of four points with respect to the fifth (one can use ρ in equation 2.5 instead of λ to avoid the change in λ values for different possible permutations). Given two sets of five points, one from the template, and the other from the input form, the algorithm computes the correspondence error as the sum of the relative errors in the five cross ratios:

$$Error = \sum_{i=1}^5 \left(\frac{\lambda_i - \Lambda_i}{\lambda_i + \Lambda_i} \right)^2 \quad (3.2)$$

where λ and Λ are from template and input form, respectively.

The five pairs of corresponding points are obtained such that equation 3.2 is minimized. This is called the Best Find strategy. The five point subset are chosen in such a way so that no three of them are collinear. Because the selected key points are located on the convex hull of the image, the order of the points is known apriori. In general, one can

use the permutation insensitive parameter ρ (equation 2.5) as an invariant signature to find the best match between two corresponding sets of five points by using equation 3.2 (replacing λ with ρ), and then verify the corresponding points using λ values. In section 3.4, this thesis presents how one can use these corresponding sets of five points to determine the affine transformation parameters required to map the input form to its corresponding template.

3.3.2 Affine Model Approach.

In this section, the distortion parameters have been modeled by an affine transformation model. Starting from the line and point features extracted from the template and input form images, one can use the following affine invariants proposed in this thesis to describe the geometry of each feature set;

- 1) Four points affine invariant $\mu=(\alpha,\beta)$.
- 2) Two parallel lines and one point affine invariant δ .
- 3) Two points and a line affine invariant γ .
- 4) Three parallel lines affine invariant η .

The above parameters are invariant under any type of affine transformation. Four different hash tables are built using the proposed invariants considering all possible permutations. The hash tables for the template are built off line and stored as a model. For each input form to be registered, one has to extract the line and point features and then build the hash tables, followed by a matching strategy to find the corresponding features between template and the input form images.

Geometric hashing strategy is used to solve the matching problem. This method is applied once for each hash table. Assume H_1 to H_4 correspond to hash tables for template and H'_1 to H'_4 correspond to hash tables for input form. Geometric hashing strategy has been applied to find the correspondence between features in both images using (H_1, H'_1) , (H_2, H'_2) , (H_3, H'_3) and (H_4, H'_4) which can be computed in parallel.

Each entity in the corresponding hash tables $(H_q, H'_q)_{q=1, \dots, 4}$ are compared with the following relative error measure.

$$\text{Error} = \frac{\|\lambda_i - \tilde{\lambda}_j\|}{\|\lambda_i + \tilde{\lambda}_j\|} \quad (3.3)$$

where λ_i and $\tilde{\lambda}_j$ come from the template and Input form Hash Tables, respectively

λ can be either μ , δ , γ or η . This algorithm uses the Euclidean Norm $\|\cdot\|$ because λ can be either a number or a vector. Assume that M_q is the resultant voting matrix corresponding to the geometric hashing between H_q and H'_q . If the template has m_p number of point features and m_r number of line features and the input form has n_p number of point features and n_r number of line features, then M_q will have m_p+m_r number of rows and n_p+n_r number of columns. The entity (k,s) (k is the index of k 'th feature in template that can be either a line or a point feature and s is the s 'th feature in input form) of the voting matrix M_q is incremented by one if their corresponding invariants λ_i and $\tilde{\lambda}_j$ has an error (equation 3.3) below a preset threshold value.

Each entity (i, j) of either one of four resultant voting matrices M_q in normalized sense represents the probability of the correspondence between feature i and j in template and input form images. Multiplying the corresponding entities in four hash tables gives the overall probability of correspondence between feature i and j in normalized sense.

$$R(i,j) = M_1(i,j) * M_2(i,j) * M_3(i,j) * M_4(i,j), \text{ for } i=1, \dots, (m_p+m_r) \text{ and } j=1, \dots, (n_p+n_r) \quad (3.4)$$

In normalized sense, $R(i,j)$ represents the probability of the correspondence between feature i from the template and feature j from the input form. The resultant matrix R has been used for the final judgment on corresponding features. In section 3.4, these corresponding results are used to determine the affine transformation parameters required to map the input form to its corresponding template.

3.4 Distortion Evaluation.

After performing the correspondence determination, using the projective model approach, five pairs of corresponding points are derived, and for affine model approach matrix R represents the probability of the correspondence between any pair of features in both template and input form.

For the first approach, using the point correspondence results $(P_i \text{ to } P'_i)_{i=1, \dots, 5}$, the transformation parameters (aff1 to aff6) are computed by solving the following set of linear equations.

$$\begin{cases} x'_i = \text{aff1} \cdot x_i + \text{aff2} \cdot y_i + \text{aff5} \\ y'_i = \text{aff3} \cdot x_i + \text{aff4} \cdot y_i + \text{aff6} \end{cases} \text{ for } i = 1, 2, \dots, 5 \quad (3.5)$$

where $P'_i = (x'_i, y'_i)^T$ in input form corresponds to $P_i = (x_i, y_i)^T$ in the template form

The number of linear equations which will be 10 are more than the number of unknowns which is 6. The best answer in least square sense would be the best least square fit to the above equations. To simplify the representation, this method uses the following matrix representations.

$$P'_i = \begin{bmatrix} \text{aff1} & \text{aff2} \\ \text{aff3} & \text{aff4} \end{bmatrix} P_i + \begin{bmatrix} \text{aff5} \\ \text{aff6} \end{bmatrix} \quad \text{or} \quad P'_i = \begin{bmatrix} \text{aff1} & \text{aff2} & \text{aff5} \\ \text{aff3} & \text{aff4} & \text{aff6} \end{bmatrix} \begin{bmatrix} P_i \\ 1 \end{bmatrix}$$

$$\text{then } [P'_1 \mid P'_2 \mid P'_3 \mid P'_4 \mid P'_5] = \begin{bmatrix} \text{aff1} & \text{aff2} & \text{aff5} \\ \text{aff3} & \text{aff4} & \text{aff6} \end{bmatrix} \begin{bmatrix} P_1 & P_2 & P_3 & P_4 & P_5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\text{if } V = [P'_1 \mid P'_2 \mid P'_3 \mid P'_4 \mid P'_5] \quad , \quad v = \begin{bmatrix} P_1 & P_2 & P_3 & P_4 & P_5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\text{then } \begin{bmatrix} \text{aff1} & \text{aff2} & \text{aff5} \\ \text{aff3} & \text{aff4} & \text{aff6} \end{bmatrix} = V v^T [v v^T]^{-1}$$

(3.6)

The matrix $v^T [v v^T]^{-1}$ is called the Pseudo Inverse Matrix of v [30]. Starting from coordinate values for these pair of five points, one can use the above equation to find the best least square fit to the transformation parameters. As a verification step, the algorithm maps the points P_1 to P_5 to the points p_1 to p_5 using the transformation parameters obtained. The transformation parameters are accepted as correct result if the following error is less than a preset threshold value.

$$\text{ERR} = \sum_{i=1}^5 \|p_i - P'_i\| \quad (3.7)$$

The input form is rejected if the above error is more than the preset threshold value.

For the second approach, this thesis uses the correspondence scores in matrix **R** resulted from the correspondence determination. Each entity (i, j) in this matrix represents the probability of the correct correspondence between feature i in the template and feature j in the input form. Features i and j can be either line or point features. The algorithm starts from the maximum entity in matrix **R** (corresponding to the most probable correspondent feature sets in both images), uses the coordinates of its corresponding feature set, and tries to estimate the transformation parameters using a modified least square fit. This modified least square fit is based on line and point correspondences. Recalling the coordinate relationship between affine transformed points and lines,

$$X = TX' \text{ where } T = \begin{bmatrix} \text{aff1} & \text{aff2} & \text{aff5} \\ \text{aff3} & \text{aff4} & \text{aff6} \\ 0 & 0 & 1 \end{bmatrix}, X = [x \ y \ 1]^T, X' = [x' \ y' \ 1]^T, L' = T^T L \quad (3.8)$$

where $L = [a \ b \ c]^T$ and $L' = [a' \ b' \ c']^T$, L and X are affine transformed to L' and X'

The best least square solution for **r** lines and **p** points correspondences would be the solution for the following equation :

$$A \cdot \begin{bmatrix} \text{aff1} \\ \text{aff2} \\ \text{aff5} \\ \text{aff3} \\ \text{aff4} \\ \text{aff6} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^p x_i x'_i + \sum_{i=1}^r a_i a'_i \\ \sum_{i=1}^p y'_i x_i + \sum_{i=1}^r a_i b'_i \\ \sum_{i=1}^p x_i + \sum_{i=1}^r a_i c'_i - \sum_{i=1}^r a_i c_i \\ \sum_{i=1}^p x'_i y_i + \sum_{i=1}^r b_i a'_i \\ \sum_{i=1}^p y_i y'_i + \sum_{i=1}^r b_i b'_i \\ \sum_{i=1}^p y_i + \sum_{i=1}^r a_i c'_i - \sum_{i=1}^r b_i c_i \end{bmatrix} = B \quad , \quad \text{where}$$

$$A = \begin{bmatrix} \sum_{i=1}^p x_i'^2 + \sum_{i=1}^r a_i'^2 & \sum_{i=1}^p x'_i y'_i & \sum_{i=1}^p x'_i & \sum_{i=1}^r a_i b_i & 0 & 0 \\ \sum_{i=1}^p x'_i y'_i & \sum_{i=1}^p y_i'^2 + \sum_{i=1}^r a_i'^2 & \sum_{i=1}^p y'_i & 0 & \sum_{i=1}^r a_i b_i & 0 \\ \sum_{i=1}^p x'_i & \sum_{i=1}^p y'_i & \sum_{i=1}^r a_i'^2 + p & 0 & 0 & \sum_{i=1}^r a_i b_i \\ \sum_{i=1}^r a_i b_i & 0 & 0 & \sum_{i=1}^p x_i'^2 + \sum_{i=1}^r b_i'^2 & \sum_{i=1}^p x'_i y'_i & \sum_{i=1}^p x'_i \\ 0 & \sum_{i=1}^r a_i b_i & 0 & \sum_{i=1}^p x'_i y'_i & \sum_{i=1}^p y_i'^2 + \sum_{i=1}^r b_i'^2 & \sum_{i=1}^p y'_i \\ 0 & 0 & \sum_{i=1}^r a_i b_i & \sum_{i=1}^p x'_i & \sum_{i=1}^p y'_i & \sum_{i=1}^r b_i'^2 + p \end{bmatrix} \quad (3.9)$$

The affine transformation parameters can be found by using the following matrix relationship.

$$\begin{bmatrix} \text{aff1} \\ \text{aff2} \\ \text{aff5} \\ \text{aff3} \\ \text{aff4} \\ \text{aff6} \end{bmatrix} = A^{-1}B \quad (3.10)$$

The matrix A can be decomposed into four 3×3 submatrices, so as to reduce the computational cost of the algorithm. The transformation parameters obtained are verified and updated using the next maximum entity of the R matrix. The final result would be the updated transformation parameters required to map the input form into its corresponding template.

In the next section, these transformation parameters are used to register or map the input form into the template.

3.5 Registration And Experimental Results.

In this section, the registration process of an input form to its corresponding template is presented. Starting from the transformation parameters obtained using the least square fit, one has to map each black pixel in the input form onto a white canvas. The resultant black and white registered input form is now aligned with the template form. Every data field inside the input form is now mapped to the same blank data field on the template form. The handwritten data can be extracted simply by addressing the corresponding location of the data field in template. The only post processing performed on the registered filled in form is the morphological closing [31] which is a Dilation

followed by an Erosion. This post processing improves the quality of the registered data in cases where the distortion consists of a serious down scaling in scanning process.

The following experimental results are presented by using a down-scaled filled-in form before and after postprocessing using morphological closing.

Figure 3.6 presents an input form which has been degraded by 75% reduction in size and some rotation. The highlighted portion of figure 3.7 is registered to the template and is presented in figure 3.7. The post processed image using a morphological closing on figure 3.7 is presented in figure 3.8.

• If you reported self-employment income on lines 135 to 143 of your return, complete this schedule.

• You have to file a statement of income and expenses with your return.

Enter the salary or wages you paid in the year to your spouse (if you are in a partnership, enter only your share). 876 25

Complete the following for the business with the largest total income:

Employer's remittance account 965-367-585

Business name united company

Location Number 465 Street 3127 MILL St P.O. Box No. NA 8 Apt. NA 8

City Winnipeg

Province or territory MB Postal code R1V 6L5

Briefly describe the major business activity:

Has the major activity changed since you filed your last return?

No ☒ Yes ☐

If the major business activity involves the resale of goods, indicate how you sell these goods: ☐ Retail ☐ Wholesale

If your major activity is trucking, are you also an "owner-operator," "leased-operator," or "broker-operator" working for a trucking concern?

No ☒ Yes ☐

Specify the principal product(s) mined, manufactured, or sold, the type of construction activity or service provided, and the approximate percentage that the revenue from each is of the total revenue.

1. 26 %

2. 65 %

3. 73 %

(Don't forget to complete Schedule 9 on the other side of this page.)

Figure 3.6 The rotated input filled-in form with 75 % scaling.

~~1ed 6000/2~~

Loc. or R.R. No.

21 3167 Mill

B.

nt

Postal code

216

business activity:

Figure 3.8 The morphological closing result of figure 3.7.

Figures 3.9a-c and figures 3.10a-c represent the template samples followed by an input form, followed by its resultant registered image. To represent the quality of registration visually, this thesis also presents figures 3.9d and figures 3.10d images obtained from the XOR operation between template and registered input form.

- If you reported self-employment income on lines 135 to 143 of your return, complete this schedule.
- You have to file a statement of income and expenses with your return.

Enter the salary or wages you paid in the year to your spouse (if you are in a partnership, enter only your share). _____

Complete the following for the business with the largest total income:

Employer's remittance account _____
 Business name _____

 Location Number, Street, P.O. Box No., or R.R. No. _____
 City _____
 Province or territory _____ Postal code _____
 Briefly describe the major business activity:

 Has the major activity changed since you filed your last return?
 No ☐ Yes ☐

If the major business activity involves the resale of goods, indicate how you sell these goods: ☐ Retail ☐ Wholesale

If your major activity is trucking, are you also an "owner-operator," "leased-operator," or "broker-operator" working for a trucking concern?

No ☐ Yes ☐

Specify the principal product(s) mined, manufactured, or sold, the type of construction activity or service provided, and the approximate percentage that the revenue from each is of the total revenue.

1. _____ %
 2. _____ %
 3. _____ %

(Don't forget to complete Schedule 9 on the other side of this page.)

Figure 3.9-a. First sample of template form.

- If you reported self-employment income on lines 135 to 143 of your return, complete this schedule.
- You have to file a statement of income and expenses with your return.

Enter the salary or wages you paid in the year to your spouse (if you are in a partnership, enter only your share). _____

356477332.00

Complete the following for the business with the largest total income:

Employer's remittance account _____
 Business name UNIV. OF WINDSOR 334-874332-0000

 Location Number, Street, P.O. Box No., or R.R. No. 660478, 3370 MILL ST. P.O. BOX 354278
 City WINDSOR
 Province or territory ONTARIO Postal code N6A 3B2
 Briefly describe the major business activity:
EDUCATIONAL AND RESEARCH PURPOSES

 Has the major activity changed since you filed your last return?
 No ☒ Yes ☐

If the major business activity involves the resale of goods, indicate how you sell these goods: ☒ Retail ☐ Wholesale

If your major activity is trucking, are you also an "owner-operator," "leased-operator," or "broker-operator" working for a trucking concern?

No ☒ Yes ☐

Specify the principal product(s) mined, manufactured, or sold, the type of construction activity or service provided, and the approximate percentage that the revenue from each is of the total revenue.

1. DATABASE AND MANUALS 35 %
 2. POWER SUPPLIES 81 %
 3. SEMINAR AND WORKSHOPS 11 %

(Don't forget to complete Schedule 9 on the other side of this page.)

Figure 3.9-b. First sample of input form.

- If you reported self-employment income on lines 135 to 143 of your return, complete this schedule.
- You have to file a statement of income and expenses with your return.

Enter the salary or wages you paid in the year to your spouse (if you are in a partnership, enter only your share)

3564777322 00

Complete the following for the business with the largest total income:

Employer's remittance account 234-874332-98600
Business name UNIV. OF WINDSOR

Location 60478, 2319 MILL ST. P.O. BOX 154278

City WINDSOR

Province or territory ONTARIO N E R 2 D 4

Briefly describe the major business activity:
EDUCATIONAL AND RESEARCH PURPOSES

Has the major activity changed since you filed your last return?

No ☒

Yes ☐

If the major business activity involves the resale of goods, indicate how you sell these goods ☒ Retail ☐ Wholesale

If your major activity is trucking, are you also an 'owner-operator,' 'leased-operator,' or 'broker-operator' working for a trucking concern?

No ☒

Yes ☐

Specify the principal product(s) mined, manufactured, or sold, the type of construction activity or service provided, and the approximate percentage that the revenue from each is of the total revenue

1	DATABOOK AND MANUALS	35	%
2	POWER SUPPLIES	81	%
3	SEMINAR AND WORKSHOPS	11	%

(Don't forget to complete Schedule 9 on the other side of this page.)

Figure 3.9-c. First sample of registered form.

3564777322 00

UNIV. OF WINDSOR 234-874332-98600

60478, 2319 MILL ST. P.O. BOX 154278

WINDSOR

ONTARIO N E R 2 D 4

EDUCATIONAL AND RESEARCH PURPOSES

DATABOOK AND MANUALS	35	%
POWER SUPPLIES	81	%
SEMINAR AND WORKSHOPS	11	%

Figure 3.9-d. Extracted information from fig. 3.9-b.

Ontario Income Tax

Ontario income tax (includes the Ontario tax reduction calculation, and the Ontario investment and employee ownership tax credit).
Guidelines appear on the back of this page. If you were not a resident of Ontario on December 31, 1993, or if you had income from a business with a permanent establishment outside Ontario in 1993, see the federal tax guide.

Section I: Ontario income tax

Basic federal tax: (Enter the amount from line 408 of your return if using the table method, or enter the amount from line 506 if using Schedule 1)			(A)
Basic Ontario tax: Calculate 58% of amount (A)			(B)
Add: Ontario surtax			
(Amount (B) minus \$5,500) X 17%	(i)		
(Amount (B) minus \$8,000) X 8%	(ii)		
Ontario surtax: Add lines (i) and (ii) above (if negative, enter "0")			(C)
Adjusted Ontario tax: Add amounts (B) and (C)			(D)
Subtract: Provincial foreign tax credit from calculation on federal Form T2036			
Ontario Income Tax			(E)

Section II: Ontario tax reduction

If you are subject to minimum tax from Form T691, or were not a resident of Ontario on December 31, 1993, you cannot claim an Ontario tax reduction.

Personal amount calculation

Basic amount		\$205	00	(F)
Amount for dependent children born in 1975 or later:				
Number of children	589	X	\$395	(G)
Amount for disabled dependants	(a)			
Number of disabled dependants	579	X	\$395	(H)
	(a)			
Total personal amount: Add amounts (F), (G), and (H).				(I)

(If amount (I) is equal to or more than amount (E), no Ontario income tax is payable. Enter "0" on line (M).)

Ontario tax reduction calculation

Enter amount from line (I)	X3		(J)
Enter amount from line (E)	X2		(K)
Ontario tax reduction: Amount (J) minus amount (K). (if negative enter "0")			(L)

Ontario tax: Amount (E) minus amount (L) (if negative, enter "0"). If you are not claiming the Ontario investment and employee ownership tax credit, enter amount (M) on line 428 of your return.

Section III: Ontario investment and employee ownership tax credit

Enter: Ontario labour-sponsored investment fund credit from OIEO 1993 certificate	541		(N)
Enter: Ontario employee ownership credit from OIEO 1993 certificate	573		(O)
Enter: Unused Ontario employee ownership credits eligible for carryforward to 1993 (see guidelines on the back of this form)			(P)
Total Ontario investment and employee ownership credit:			
Add amounts (N), (O), and (P)			(R)
Ontario tax: Amount (M) minus amount (R) (if negative, enter "0")			
Enter this amount on line 428 of your return.			(S)

Figure 3.10-a : Second sample of template form.

Ontario Income Tax

Ontario income tax (includes the Ontario tax reduction calculation, and the Ontario investment and employee ownership tax credit).
Guidelines appear on the back of this page. If you were not a resident of Ontario on December 31, 1993, or if you had income from a business with a permanent establishment outside Ontario in 1993, see the federal tax guide.

Section I: Ontario income tax

Basic federal tax:	(Enter the amount from line 406 of your return if using the table method, or enter the amount from line 506 if using Schedule 1)	76565378.75 (A)	37676631.00 (B)
Basic Ontario tax:	Calculate 56% of amount (A)	6564578.00	47.86
Add: Ontario surtax:	(Amount (B) minus \$5,500) x 17% (A) 32236262.37	3523565.3672	627.53
Ontario surtax:	Add lines (I) and (J) above (if negative, enter "0")	3672	36357637.33
Adjusted Ontario tax:	Add amounts (B) and (C)		32878272.00 (C)
Subtract: Provincial foreign tax credit from calculation on federal Form T2096			83787222.88 (C)
Ontario income tax:			22238273.11 (E)

Section II: Ontario tax reduction

If you are subject to minimum tax from Form T661, or were not a resident of Ontario on December 31, 1993, you cannot claim an Ontario tax reduction.

Personal amount:		\$806.00 (F)
Amount for dependent children born in 1975 or later:	5 x \$395	36535.22 (G)
Amount for disabled dependents:	2 x \$395	43763.11 (H)
Number of disabled dependents:	0	3476.55 (I)
Total personal amount:	Add amounts (F), (G), and (H)	37827387.22 (J)
Ontario tax reduction calculation:	(If amount (I) is equal to or more than amount (E), no Ontario income tax is payable. Enter "0" on line (M).)	54884388.00 (K)
Ontario tax reduction:	Amount (J) minus amount (K). (If negative, enter "0")	8438444.54 (L)

Section III: Ontario investment and employee ownership tax credit

Ontario tax:	Amount (E) minus amount (L) (if negative, enter "0")	32878222.13
Ontario investment and employee ownership tax credit:	Enter amount from line (E)	47663833.33 (M)
Ontario tax:	Amount (E) minus amount (L) (if negative, enter "0")	32878222.13
Ontario investment and employee ownership tax credit:	Enter amount from line (E)	47663833.33 (M)
Ontario tax:	Amount (E) minus amount (L) (if negative, enter "0")	32878222.13
Ontario investment and employee ownership tax credit:	Enter amount from line (E)	47663833.33 (M)
Ontario tax:	Amount (E) minus amount (L) (if negative, enter "0")	32878222.13
Ontario investment and employee ownership tax credit:	Enter amount from line (E)	47663833.33 (M)
Ontario tax:	Amount (E) minus amount (L) (if negative, enter "0")	32878222.13
Ontario investment and employee ownership tax credit:	Enter amount from line (E)	47663833.33 (M)

Figure 3.10-b : Second sample of Input form.

Ontario Income Tax

Ontario income tax (includes the Ontario tax reduction calculation, and the Ontario investment and employee ownership tax credit).
Guidelines appear on the back of this page. If you were not a resident of Ontario on December 31, 1993, or if you had income from a business with a permanent establishment outside Ontario in 1993, see the federal tax guide.

Section I: Ontario income tax

Basic federal tax: (Enter the amount from line 406 of your return if using the table method, or enter the amount from line 506 if using Schedule 1)		6564578.00	76565272.75 (A)	
Basic Ontario tax: Calculate 55% of amount (A)		47.86	2767611.75	437676631.00 (B)
Add: Ontario surtax				
(Amount (B))	12	minus \$5,500 X 17%	3523565	(i) 32236262.37
(Amount (B))	11	minus \$8,000 X 8%	3672	(ii) 627.53
Ontario surtax: Add lines (i) and (ii) above (if negative, enter "0")			36357637.33	32878787.00 (C)
			Adjusted Ontario tax: Add amounts (B) and (C)	
			83787222.89	(D)
Subtract: Provincial foreign tax credit from calculation on federal Form T2036				
			82398822.99	
			27238782.11	(E)
			Ontario Income Tax	

Section II: Ontario tax reduction

If you are subject to minimum tax from Form T691, or were not a resident of Ontario on December 31, 1993, you cannot claim an Ontario tax reduction.

Personal amount calculation

Basic amount		\$205	00 (F)
Amount for dependent children born in 1975 or later:			
Number of children	5	x \$395	36535.22 (G)
Amount for disabled dependants:			
Number of disabled dependants	570	2 x \$395	43763.11 (H)
Total personal amount: Add amounts (F), (G), and (H).		3476	55 (I)

(If amount (I) is equal to or more than amount (E), no Ontario income tax is payable. Enter "0" on line (M).)

Ontario tax reduction calculation

Enter amount from line (I)	32877822.22	x3	37827387.22 (J)
Enter amount from line (E)	47663833.33	x2	54884888.00 (K)
Ontario tax reduction: Amount (J) minus amount (K). (If negative enter "0")			84384844.54 (L)

Ontario tax: Amount (E) minus amount (L) (if negative, enter "0"). If you are not claiming the Ontario investment and employee ownership tax credit, enter amount (M) on line 428 of your return.

374734.00 (M)

Section III: Ontario investment and employee ownership tax credit

Enter: Ontario labour-sponsored investment fund credit from OIEO 1993 certificate	541	43988844.00 (N)
Enter: Ontario employee ownership credit from OIEO 1993 certificate	573	43874874.88 (O)
Enter: Unused Ontario employee ownership credits eligible for carryforward to 1993 (see guidelines on the back of this form)		53984844.99 (P)
Total Ontario investment and employee ownership credit:		48983833.00
Add amounts (N), (O), and (P):		477433.00 (R)
Ontario tax: Amount (M) minus amount (R) (if negative, enter "0")		4783787833
Enter this amount on line 428 of your return.		8788437.99 (S)

Figure 3.10-c : Second sample of registered form.

	6564578.00	7656527; 75	
	47.86	2767611 75	43767663; 00
12	3523565	3223626; 37	
11	3672	62; 53	
		3635763; 33	32878787 00
			83787222 09
			82388822 09
			27238783 1
	5	36535 22	
	2	43763 11	
		3476 55	
	32877822 22	37827387 22	
	47663833 33	54884388 00	
		84384844 54	478873 09
			374734 00
		43989844 0	
		43874874 8	
		53984844 99	
		489838331 0	477433 00
		4783787833	8788437 09

Figure 3.10-d : Extracted information from 3.10-b.

3.6 Summary and Concluding remarks.

The problem of registering a filled-in form into a machine readable form for information extraction is a challenging research area in Pattern Recognition and Document Processing. In this chapter, a new form registration approach using the results from computer vision is presented. The geometric distortions that occur in scanned input filled-in form have been modeled by an affine transformation. The algorithm is paying the trade-off of using a blank sample of the form by covering any global distortion that may happen to the input form. This advantage makes this approach robust enough to accept non-uniformly scaled, sheared, rotated, translated and even reflected input forms compare to previous approaches which just cover the rotation only distortions. The template relevant features are extracted and their geometric relations have been modeled by different projective and affine signatures off-line. These affine invariants were used for matching between different relevant features in template and input filled-in form. The input form relevant features have been extracted and their geometric relations have been obtained on-line using the same distortion model. Two different matching strategies have been used for correspondence determination. The Best-Find strategy performed weaker than the geometric hashing approach, since there is a possibility of occlusion and addition of relevant features due to noise. The geometric hashing approach has been even improved further by using a voting integration based on statistical point of view. The experimental studies have shown that this method is promising and powerful enough to cover any global distortions happening in input forms.

In the next chapter, the sensitivity analysis and OCR performance evaluation of the proposed method are presented. In addition to the previous simulation results, extensive tests on various kinds of forms under variety of distortions have been carried out. At the end of this thesis, we have presented a number of these results. These images are the template blank forms followed by the filled in form with the geometric distortion present, and the extracted information after the application of the proposed form registration method.

CHAPTER IV

Performance Evaluation

This chapter deals with sensitivity analysis and OCR performance for evaluation of the proposed method. Performance is evaluated for documents that contain noise and different types of distortions. This sensitivity analysis and performance evaluation of the proposed method will be carried out in each of the following steps involved in the proposed algorithm;

- 1) Feature Extraction.
- 2) Affine Invariant based geometric modeling.
- 3) Geometric hashing.
- 4) Registration and post-processing.

The performance of the feature extraction step is evaluated by presenting results on the robustness of the feature extraction in noisy documents. Impulsive noise is added to the input filled-in form.

Sensitivity analysis of the invariants has been carried out when the relevant features are corrupted by noise with a normal distribution, zero mean and a known standard deviation σ .

In order to test the robustness of the algorithm based on different kinds of distortions available, we present simulation results using different affine transformed filled-in forms. The geometric distortion in input filled-in forms are simulated by applying

different primitive distortions to aligned filled-in forms such as rotation, scaling, and shearing.

4.1 Sensitivity Analysis.

The features used are the keypoints on the convex hull and line features in input images. Pepper noise is added to the image of a sample form. Line and point features are extracted from the noisy image and compared to the true noise free features. The mean and standard deviation of the error are evaluated as a function of the level of noise added.

4.1.1 Sensitivity Analysis of the feature extraction.

The form registration method uses the convex hull edges of the black and white image of the input form as point features if they belong to a connected component of at least 16 black pixels. Line features are extracted using a modified Hough transform approach that includes a connected component analysis and a line fitting algorithm. To measure the sensitivity of the feature extraction, pepper noise is added in proportion to the number of black pixels in the image. If, for example, an input form has N black pixels, adding 100% noise to the document implies adding N black pixels in random locations of the document. The locations where the noise pixels are to be added are generated using a random number generator. A portion of an input form and its corresponding noisy image (100% noise added) are shown in figures 4.1 and 4.2.

The test consists of adding noise in steps of 10% from 0 to 200% to a selected form. At each step the features are extracted and compared to the true features extracted from the noise free form. The average distance is computed in each test.

In the case of point features, the algorithm starts to miss one keypoint out of 8 keypoints selected as point features when the level of noise added exceeds 90%.

Postal code	Sp
	typ
	apl
	rev
	1.
your last return?	2.
	3.
s ☐	

Figure 4.1. A portion of the input form.

Postal code	Sp
	typ
	apl
	rev
	1.
your last return?	2.
	3.
s ☐	

Figure 4.2. 100% noise added to Figure 4.1.

The worst case was an error of 6 pixels in extracted point feature when the noise added is more than 90%. The average of the error in point coordinates is shown in the figure 4.3.

Figure 4.3. Average error in keypoints extracted due to noise added to the form.

In the case of line features, there were no errors for noise levels up to 200%. The average and standard deviation of the distance between noise free line features and noisy line features are computed in each test. It was noted that the errors were negligibly small.

The error in detected lines is calculated based on the absolute error in slope of the line and the relative error in y-axis intersection of the line. As expected, the line detection is more robust with regard to keypoint extraction. Even if there is an error in extracting the relevant features, the overall approach relies on the correspondence

determination between features in template and input form starting from the best corresponding features and there is always a verification and updating step which helps to avoid selecting false correspondences. The connected component analysis of the keypoint feature extraction helps to avoid noisy blobs as relevant points. The keypoint features added or occluded due to noise give rise to a number of spurious invariants in the input form hash tables. These false signatures are discarded in the correspondence determination, since there are no corresponding signatures in template hash tables to confirm a match. Experimental studies with different input form layouts have shown the reliability of the proposed method in a noisy environment. Since the error in detecting lines is almost zero, this thesis continues with the sensitivity analysis of the four point affine invariant only in the next section.

4.1.2 Sensitivity Analysis of the affine invariants.

Line features can be extracted with much smaller error compared to the keypoint features that are sensitive to pepper noise in the document. This section provides test results on the four points $\mu=(\alpha,\beta)$ invariant in order to prove the insensitivity of the invariant used. The other three affine invariants show negligible changes to the error in keypoints because their computation involves line features.

Here we use the following linear approximation to analyze the behaviour of $\mu=(\alpha,\beta)$ to changes in point coordinates.

$$\Delta\alpha = \sum_{i=1}^4 \frac{\partial\alpha}{\partial x_i} \Delta x_i + \sum_{i=1}^4 \frac{\partial\alpha}{\partial y_i} \Delta y_i, \quad \Delta\beta = \sum_{i=1}^4 \frac{\partial\beta}{\partial x_i} \Delta x_i + \sum_{i=1}^4 \frac{\partial\beta}{\partial y_i} \Delta y_i \quad (4.1)$$

Assuming independent normal distributions for the error in point coordinates with zero mean and standard deviation σ ($N(0,\sigma)$), one can show that both $\Delta\alpha$ and $\Delta\beta$ have zero mean and their standard deviations are proportional to the absolute values of α and β . The statistical behavior of $\mu=(\alpha,\beta)$ for different geometries of sets of four points is presented using equation (4.1) (here we assume independent normal distributions for error in point coordinates with zero mean and standard deviation σ ($N(0,\sigma)$), to show the sensitivity of $\mu=(\alpha,\beta)$ to errors in point coordinates. To measure the sensitivity of $\mu=(\alpha,\beta)$, one can carry out the following derivations to find the standard deviation of $\Delta\alpha$, $\Delta\beta$ in terms of a set of four point coordinates assuming a linear approximation.

$$\begin{aligned}\sigma_{\Delta\alpha}^2 &= \sigma^2 \cdot \left(\sum_{i=1}^4 \left(\frac{\partial\alpha}{\partial x_i} \right)^2 + \sum_{i=1}^4 \left(\frac{\partial\alpha}{\partial y_i} \right)^2 \right) = C_{\alpha} \cdot \sigma^2, \\ \sigma_{\Delta\beta}^2 &= \sigma^2 \cdot \left(\sum_{i=1}^4 \left(\frac{\partial\beta}{\partial x_i} \right)^2 + \sum_{i=1}^4 \left(\frac{\partial\beta}{\partial y_i} \right)^2 \right) = C_{\beta} \cdot \sigma^2, \\ \text{cov}(\Delta\alpha, \Delta\beta) &= \sigma^2 \cdot \left(\sum_{i=1}^4 \left(\frac{\partial\alpha}{\partial x_i} \right) \left(\frac{\partial\beta}{\partial x_i} \right) + \sum_{i=1}^4 \left(\frac{\partial\alpha}{\partial y_i} \right) \left(\frac{\partial\beta}{\partial y_i} \right) \right) = C_{\alpha\beta} \cdot \sigma^2\end{aligned}\tag{4.2}$$

MAPLE software was used to compute the differentiations which appear in equation (4.2). The values of C_{α} , C_{β} and $C_{\alpha\beta}$ were evaluated for different geometries of sets of four points, no three of them collinear and their coordinates normalized to a square centered at the origin and 1000 pixels wide. The average of the followings :

- $\left(\frac{\sqrt{C_{\alpha}}}{\alpha} \times 100 \% \right)$

- $\left(\frac{\sqrt{C_\beta}}{\beta} \times 100 \% \right)$
- $\left(\sqrt{\frac{C_{\alpha\beta}}{\alpha\beta}} \times 100 \% \right)$

which are good measures of the sensitivity of α and β to different sets of four points, were 0.8643%, 0.9925% and 0.7003% which are less than 1%. As a result one can find :

- $\sigma^2 \Delta\alpha < 10^{-4} \alpha^2 \sigma^2$
- $\sigma^2 \Delta\beta < 10^{-4} \beta^2 \sigma^2$
- $\text{Cov}(\Delta\alpha, \Delta\beta) < 10^{-4} \alpha\beta \sigma^2$

The results show that $\mu=(\alpha,\beta)$ is well-behaved to changes in point coordinates. A closed form expression for C_α , C_β and $C_{\alpha\beta}$ are too cumbersome to include in this thesis.

4.1.3 Robustness.

It can be shown that an affine transformation can be represented as the multiplication of an x-shearing matrix, anisotropic scaling matrix and a rotation matrix which makes $\mathbf{M}$ plus a translation vector $\mathbf{d}$ in equation (4.3) compared to equation (2.6)[16].

$$\mathbf{M} = \begin{bmatrix} \text{aff}_1 & \text{aff}_2 \\ \text{aff}_3 & \text{aff}_4 \end{bmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \quad \mathbf{d} = \begin{bmatrix} \text{aff}_5 \\ \text{aff}_6 \end{bmatrix} \quad (4.3)$$

Since the translation vector $\mathbf{d}$ has just the effect of shifting the image in both horizontal and vertical directions, the geometric distortion is normally represented by the

matrix M . To see the overall performance of the proposed method in the presence of global affine distortion, this thesis simulates different primitive distortions using the parameters a, b, c and ϕ in equation 4.3. The proposed algorithm is applied to the simulated forms and the distortion parameters are extracted. The performance is presented by introducing the difference between the true geometric distortion and evaluated distortion parameters defined in equation 4.4;

$$ERR = \|M - \bar{M}\| = \sqrt{\sum_{i=1}^4 (aff_i - \overline{aff_i})^2} \quad (4.4)$$

where M represents the true distortion parameters and $\bar{M}$ represents the distortion evaluated by applying the algorithm to the input form.

Tests were carried out for four different forms with different sizes (from 2500×1700 to 2500×3500 pixels in 300 dpi scanning mode) and different layout structures. Figures 4.4 to 4.7 represent the four template forms used in this test.

- If you reported self-employment income on lines 135 to 143 of your return, complete this schedule.
- You have to file a statement of income and expenses with your return.

Enter the salary or wages you paid in the year to your spouse (if you are in a partnership, enter only your share). _____

Complete the following for the business with the largest total income:

Employer's remittance account _____
Business name _____

Location Number, Street, P.O. Box No. or R.R. No. _____
City _____
Province or territory _____ Postal code _____

Briefly describe the major business activity: _____

Has the major activity changed since you filed your last return?
No ☐ Yes ☐

If the major business activity involves the resale of goods, indicate how you sell these goods: ☐ Retail ☐ Wholesale

If your major activity is trucking, are you also an "owner-operator," "leased-operator," or "broker-operator" working for a trucking concern?

No ☐ Yes ☐

Specify the principal product(s) mined, manufactured, or sold, the type of construction activity or service provided, and the approximate percentage that the revenue from each is of the total revenue.

1. _____ %
2. _____ %
3. _____ %

(Don't forget to complete Schedule 9 on the other side of this page.)

Figure 4.4. First form used for the robustness test.

Ontario Income Tax

Ontario income tax (includes the Ontario tax reduction calculation, and the Ontario investment and employee ownership tax credit).
Guidelines appear on the back of this page. If you were not a resident of Ontario on December 31, 1993, or if you had income from a business with a permanent establishment outside Ontario in 1993, see the federal tax guide.

Section I: Ontario income tax

Basic federal tax: (Enter the amount from line 406 of your return if using the table method, or enter the amount from line 506 if using Schedule 1)

Basic Ontario tax: Calculate 56% of amount (A)		(A)	(B)
Add: Ontario surtax			
(Amount (B) minus \$5,500) X 17%	(i)		
(Amount (B) minus \$8,000) X 8%	(ii)		
Ontario surtax: Add lines (i) and (ii) above (if negative, enter "0")			(C)
Adjusted Ontario tax: Add amounts (B) and (C)			(D)
Subtract: Provincial foreign tax credit from calculation on federal Form T2036			
Ontario Income Tax			(E)

Section II: Ontario tax reduction

If you are subject to minimum tax from Form T691, or were not a resident of Ontario on December 31, 1993, you cannot claim an Ontario tax reduction.

Personal amount calculation

Basic amount	\$205	00	(F)
Amount for dependent children born in 1975 or later:			
Number of children	569	x \$395	(G)
Amount for disabled dependants:			
Number of disabled dependants	570	x \$395	(H)
Total personal amount: Add amounts (F), (G), and (H).			(I)

(If amount (I) is equal to or more than amount (E), no Ontario income tax is payable. Enter "0" on line (M).)

Ontario tax reduction calculation

Enter amount from line (I)	x3	(J)
Enter amount from line (E)	x2	(K)
Ontario tax reduction: Amount (J) minus amount (K). (if negative enter "0")		(L)
Ontario tax: Amount (E) minus amount (L) (if negative, enter "0"). If you are not claiming the Ontario investment and employee ownership tax credit, enter amount (M) on line 428 of your return.		(M)

Section III: Ontario investment and employee ownership tax credit

Enter: Ontario labour-sponsored investment fund credit from OIEO 1993 certificate	541	(N)
Enter: Ontario employee ownership credit from OIEO 1993 certificate	573	(O)
Enter: Unused Ontario employee ownership credits eligible for carryforward to 1993 (see guidelines on the back of this form)		(P)
Total Ontario investment and employee ownership credit: Add amounts (N), (O), and (P)		(R)
Ontario tax: Amount (M) minus amount (R) (if negative, enter "0")		(S)
Enter this amount on line 428 of your return.		

Figure 4.5. Second form used for the robustness test.

T1-1993
Federal tax

Detailed Tax Calculation — Schedule 1

1993 rates of federal income tax

Taxable income	Tax rate
less than \$29,590	17%
\$29,590 to \$59,180	\$ 5,030 on the first \$29,590, plus 26% of the taxable income over \$29,590
more than \$59,180	\$12,724 on the first \$59,180, plus 29% of the taxable income over \$59,180

Federal income tax — Use the "1993 rates of federal income tax" above.

Taxable income from line 400 of your return

On the first	_____	tax is	_____
On the remaining	_____	tax at	% is _____

Total federal income tax on taxable income

Add: Tax adjustments (see line 500 in the guide) Specify:

500

Subtract: Total non-refundable tax credits from line 350 of your return

501

Federal dividend tax credit: 13.33% of taxable amount of dividends from taxable Canadian corporations (line 120 of your return)

502

Minimum tax carry-over (see line 504 in the guide)

504

Total of above credits

Basic federal tax 506

Subtract: Federal foreign tax credit (make a separate calculation for each foreign country)

503

Income tax or profits tax paid to a foreign country (see line 507 in the guide)

507

net foreign income * 508

net income **

X ("Basic federal tax" plus any dividend tax credit) =

- * You have to reduce your net foreign income by any foreign income exempt under a tax treaty (included on line 256).
- ** Net income (line 236) (or if you filed a Form T581 election, use line 7 of that form; if negative, enter "0") minus any capital losses of other years allowed (line 253), employee home relocation loan deduction (line 248), stock option and shares deductions (line 249), other payments deduction (line 250), capital gains deduction (line 254), and any foreign income exempt under a tax treaty (included on line 256).
- *** You have to reduce your basic federal tax by any tax adjustments for CPP/QPP disability benefits for previous years (line 500).

Deduct (a) or (b) above, whichever is less 509

Federal tax (enter this amount on line 406 of your return) 406

Federal individual surtax (see line 419 in the guide)

Basic federal tax (line 506)

Subtract: Federal forward-averaging tax credit (attach a completed Form T581)

Amount (A)

Amount (A)

x 3%

(Amount (A))

minus \$12,500) x 5%

Individual surtax (add lines (i) and (ii) above)

Subtract: Additional federal foreign tax credit from Part II of Form T2209

510

511

Subtotal

Subtract: Additional investment tax credit from section II of Form T2036 (IND.)

517

518

Federal individual surtax (enter this amount on line 419 of your return) 419

Ontario income tax

Use Form T1C (ONT.) TC-1993 to calculate your Ontario tax.

Figure 4.6. Third form used for the robustness test.

Direct Deposit Request Form for Individuals

- Complete this form to have your income tax refund, goods and services tax (GST) credit, or child tax benefit payments deposited directly.
- Attach the completed form to your income tax return or send it to the nearest Revenue Canada income tax office.
- For more information on Direct Deposit, see the back of this form.

Part A – Identification (Please print clearly)		
First name and initial	Last name	Social insurance number
Address		Is this a new address? Yes ☐ No ☐
City	Province or territory	Postal code
Part B – To start Direct Deposit or change information		
<p>► To have your income tax refund and GST credit deposited directly, enter the banking information in areas 702, 703 and 704:</p>		
702 Branch number	703 Institution number	704 Account number
Give name(s) of account holder(s), if different from Part A:		
<p>► To have your child tax benefit deposited directly, check box 701 OR enter the banking information in areas 705, 706 and 707:</p>		
☐ 701 I want my benefit deposited in the same account as above.		
705 Branch number	706 Institution number	707 Account number
Give name(s) of account holder(s), if different from Part A:		
Part C – To confirm banking information		
Attach a personalized cheque with "VOID" written on the front, OR have your financial institution complete Part C.		
Name, address and postal code of financial institution (A stamp can be used)		
Signature of financial institution official		
Telephone no. of financial institution		Date Year Month Day
Name, address and postal code of financial institution (A stamp can be used)		
Signature of financial institution official		
Telephone no. of financial institution		Date Year Month Day
Part D – To stop Direct Deposit		
☐ 708 Indicate for which payment(s) you wish to stop Direct Deposit (Check one box only):		
1 ☐ All payments 2 ☐ Income tax refund and GST credit 3 ☐ Child tax benefit		
Part E – Authorization		
I, as the person entitled to receive the above-mentioned payment(s) and in lieu of my receiving a cheque for the same from the Receiver General, hereby authorize the Receiver General for Canada to deposit, until further notice, the payment(s) described above into the account(s) noted herein by means of Direct Deposit.		
I also agree that if Revenue Canada, Customs, Excise and Taxation is advised by my financial institution of a change in my financial institution number, branch number or account number, it may redirect payments to me in accordance with the information provided by my financial institution.		
Signature		Date Year Month Day
Telephone number		Date

Figure 4.7. Fourth form used for the robustness test.

Each form has been distorted by primitive distortions using the distortion model in equation 4.3. The rotation only parameter, ϕ , was changed in 10-degree steps from zero to 360 degrees. The shearing only parameter, b , was varied in 0.1 steps from zero to 0.7. The scaling parameters assumed to be equal ($a=c$) and varied between 0.7 to 1.3 in 0.1 steps. Approximately 1000 tests were carried out to evaluate the algorithm. The following table (4.1) shows the average and the worst case errors in transformation parameters (Error in equation 4.4) obtained as a result of applying this algorithm compared to the true transformation parameters to different forms under test.

Average Error	31.332×10^{-4}	4.049×10^{-4}	4.805×10^{-4}
Worst Case Error	107.4×10^{-4}	16.56×10^{-4}	13.38×10^{-4}

Table 4.1 : The average and the worst case errors obtained in transformation parameters.

To clarify the results obtained in table 4.1, assume there is a 2000 by 2000 pixels wide square that is to be registered using the proposed method. The average distance error between the vertices of the registered square and the true position of these vertices are computed using the error presented in table 4.1. The average and the worst case scenario of this distance error for each type of distortions involved are measured and presented in the following table 4.2.

Average Distance	5.5	0.68	0.34
Worst Case Distance	17	2.55	1.15

Table 4.2 : Average and worst case errors obtained with form documents.

Experimental studies have shown that the proposed algorithm is robust to different types of distortion. Table 4.2 displays the average and worst case errors obtained over a class of form documents. It is seen that the errors are within the acceptable range considering the size of our images.

4.1.4 Sensitivity Analysis Remarks

The results of the sensitivity analysis of the features show that the line features can be extracted with relatively smaller error compared to the errors associated with the extraction of keypoints from noisy images. However, the errors associated with the keypoint features are also within reasonable limits. The use of invariants based on both point and line features increases the reliability of the form registration process when compared invariant matching strategies using only keypoint features.

The study of sensitivity analysis of the affine invariants were focused only on the four points affine invariant, since it is the only invariant which is computed based on key points alone. The relationship between the statistical behavior of this invariant and the noise model for the error in estimation of point coordinates have been evaluated. This study has shown that the proposed invariants are relatively insensitive to the errors in the estimation of the features.

The overall performance of the proposed algorithm in the presence of noise and different geometric distortions has proven to be satisfactory and promising.

4.2 OCR evaluation.

Information retrieval and storage are the final stages in any document processing system. The filled-in information inside the registered form should be examined by an Optical Character Recognition (OCR) system to be recognized and stored properly. Considering the importance of the outcome of this stage, one has to evaluate the performance of the registration method based on the OCR results. Any document registration method degrades the quality of the registered data to some extent. The post-processing stage is normally used to correct these degradations. The registration approach chosen should be required to preserve the quality of the filled-in information. In this section, the quality of the registered forms are evaluated and compared to the OCR results obtained from non-distorted forms.

the next step in OCR evaluation of the proposed method is to use OCR techniques to evaluate the performance of the non-registered form and the registered. The improvement, if any, will clearly bring out the advantages of the proposed registration techniques.

4.2.1 OCR Evaluation of the Registered Data.

In form registration systems, the quality of the registered data is as important as the registration itself since the final recognition stage relies on the quality of the registered

data. One can measure the quality of the registered data using a commercial OCR package to read the data inside the registered form. The comparison is performed between the OCR results on the non-distorted form and registered form for different types of distortion. Tests were carried out on four different forms with different sizes and different layout structures as shown in Figures 4.4 to 4.7.

Each form has been distorted by primitive distortions using the distortion model in equation 4.3. The rotation parameter ϕ was changed in 10-degree increments from zero to 90 degrees. The shearing parameter b was varied in increments of 0.1 from zero to 0.7. The scaling parameters assumed to be equal ($a=c$) were varied between 0.7 to 1.3 in steps of 0.1. Each distorted input form is registered using the proposed registration method. A post-processing morphological closing operation was used to increase the quality of the registered data in cases where serious down-scaling was present. *TypeReader OCR* software was used for recognition of text in the form. The resulting recognition rate is computed for each registered image and compared to its corresponding template form recognition rate. The average and the standard deviation of the difference in recognition results for each form and each type of distortion are presented in table 4.3.

Average (%)	1.3	2.7	1.6
Variance	0.91	0.75	0.76

Table 4.3 : Average and variance of the reduction in recognition rate after application of the proposed algorithm compared to the non-distorted templates.

This table presents the reduction in recognition rate resulting from the degradation made by using the proposed form registration algorithm. It is evident that the reduction in accuracy is relatively small.

4.2.2 Overall Performance Evaluation of the Algorithm.

Considering the importance of the quality of the registered data, this thesis presents some test results on the overall performance of the proposed method. The previous OCR package (*TypeReader Professional*) has been used to carry out these tests. In these tests, the performance of the proposed method is evaluated by measuring the improvement in recognition rate made by the proposed method compared to the recognition results on non-registered forms. Four different forms have been used for these tests. Each form has been distorted by primitive distortions using the distortion model in equation 4.3. The rotation parameter ϕ was changed in 10-degree increments from zero to 90 degrees. The shearing parameter b was varied in increments of 0.1 from zero to 0.7. The scaling parameters assumed to be equal ($a=c$) was varied between 0.7 to 1.3 in 0.1 steps. Each distorted input form is registered using the proposed registration method.

Each distorted form was processed by the OCR software to evaluate the performance. The graphs shown in figures 4.8 to figure 4.10 present the average recognition rate for non-registered forms in the presence of each geometric distortion. The averaging was performed on four different test forms. The results show the poor performance of the OCR package on geometric distorted input documents.

Figure 4-8 : Average recognition rate of the non-registered forms in the presence of shear.

Figure 4-9 : Average recognition rate of the non-registered forms as a function of scaling.

Figure 4-10 : Average recognition rate of the non-registered forms in the presence of rotation.

The distorted forms are registered using the proposed form registration method and the registered form was processed by the same OCR software. The average recognition rates using four different forms are displayed in figures 4.11 to 4.13.

Figure 4-11 : Average recognition rate of the registered forms as a function of shearing.

Figure 4-12 : Average recognition rate of the registered forms as a function of scaling.

Figure 4-13 : Average recognition rate of the registered forms as a function of rotation.

The overall performance of the proposed method can be demonstrated by evaluating the improvement in recognition rate made by using the proposed method. This can be achieved by subtracting the recognition rate of the non-registered forms from the recognition rate of registered forms in the presence of each type of distortion. The results are presented in figures 4.14 to 4.16. These graphs display the average improvement made by using the proposed registration method to register the geometric distorted forms.

Figure 4-14 : Average improvement in recognition rate of the registered forms considering the effect of shearing.

Figure 4-15 : Average improvement in recognition rate of the registered forms in the presence of scaling.

Figure 4-16 : Average improvement in recognition rate of the registered forms considering the effect of rotation.

4.2.3 Conclusions.

Following the sensitivity analysis, the evaluation of the performance of the proposed algorithm was undertaken by considering the final OCR results on processed forms. The OCR package used show very poor performance dealing with non-registered forms when there is a geometric distortion present(see to figures 4.8 to 4.10). After the application of the proposed method to register the input forms, the OCR results show considerable improvement (see to figures 4.11 to 4.13). Figures 4.14 to 4.16 display the overall improvement obtained using the proposed registration method. The quality of the registered data is degraded and the average reduction in recognition rate has been presented in table 4.3. This amount of degradation is negligible, considering the overall improvement made by applying the proposed method to correctly recognize the geometric distorted forms that are normally being rejected by commercial OCR packages. The

overall performance evaluation results have proven the proposed method to be robust to different geometric distortions present in input forms.

4.3 Discussions.

The performance of the feature extraction algorithm has proven its robustness to noisy documents. The line detection performed better than the keypoint extraction, since lines are more tolerant to noise than keypoints.

The statistics of the error in affine invariants used are obtained based on a normal distribution assumption in error in keypoint features with zero mean and standard deviation σ . The results show the robust behavior of the invariants chosen to error in relevant features extracted from the image.

The robustness of the overall registration algorithm to different geometric distortions has been evaluated. The error between the true distortion parameters and the evaluated distortion parameters have been computed and presented in table 4.1. Based on these results, another interpretation based on the distance error on a 2000 by 2000 pixel wide square has been introduced to clarify the previous results. Experimental studies have shown that the proposed algorithm is robust to different types of distortion and the average and worst case errors are within the acceptable range considering the size of our images. The overall performance of the proposed method based on an OCR point of view has been tested. The improvement in recognition rate has been evaluated and presented in figures 4-14 to 4-16.

CHAPTER V

Summary and Conclusions

5.1 Summary.

The main objective of this dissertation is to develop new techniques to facilitate machine processing of form documents. Research is focused on form registration which is a major subtask of any form processing system. The goal is to correct any geometric distortion normally found in scanned forms. An important criterion was the removal of any restriction on the level of distortion in the form to be processed. The only requirement, however, is the availability of a blank sample of the form to be processed.

The geometric distortion in scanned form normally consists of rotation, translation, (non-uniform) scaling, shearing, or any combination of these primitive distortions that can be modeled as an affine transformation in a plane. A blank sample of the form which is suitably scanned (no geometric distortion) and stored as the model to represent its geometric properties in a suitable format. Using this representation of the model, one can perform the same geometric modeling of the filled-in form that has undergone distortion. The matching between these two representations gives the corresponding geometric features in both input form and the model form. Using these corresponding features, the transformation parameters required to map the input filled-in form to its corresponding non-distorted template are estimated. The result of this mapping is used to obtain the distortion corrected filled-in form which is aligned and ready to be processed. Each filled-

in field of the input form is now mapped to its corresponding field in blank sample of the form.

The motivation for this approach comes from a computer vision object recognition strategy. Using invariant geometric modeling, the goal is to find the corresponding relevant features between two affine transformed images (an object to be recognized and the model image). The geometric features used are line and key-point features. The straight lines inside the form and the Convex Hull subset of the form have proven to be the most suitable choices for relevant features. The geometric model representation is found using a collection of affine invariants, calculated by using the extracted features. The matching is performed using a geometric hashing voting scheme. The voting results are integrated based on an statistical model. A least square approach is used to find the optimum transformation parameters required to correct the distortion in filled-in form. The proposed algorithm has proven to be feasible, reliable and robust to noisy documents and different geometric distortions. Experimental studies on real and simulated forms have shown the new proposed form registration method to be promising.

5.2 Conclusions.

In the area of form processing, this is the first time that the problem of form registration has been tackled in such a general form. Various models for the registration of forms that have undergone an affine distortion have been developed. These contributions can be summarized as follows :

- Solving the registration problem using projective invariants.

- Solving the registration problem using affine invariants.
- Integrating four different affine invariants to solve the registration problem
- Using a geometric hashing technique followed by a statistical integration of voting results to improve the extraction of the geometric distortion.
- Removal of any restriction on the magnitude of the different types of distortion.

5.3 Future Research.

In this section, some suggestions are provided on pursuing the proposed research and introduces the topics that are worthy of future investigation. The algorithm relies on finding the correspondence between the relevant features on the input form and template form.

Since the location of the key points has a significant impact on the accuracy of the registration, one might investigate the optimal locations for placing the relevant features (lines and/or landmarks) inside the form. Geometric feature sets can be placed in specific locations on the form to enable one to improve the form registration performance even further.

Next issue which is worthy of further investigation would be using the statistical properties of the entities of the template hash table, computed based on each set of extracted relevant feature, to improve the geometric hashing strategy using a weighted voting scheme. This modification will increase the robustness of the matching strategy by considering the behavior of different geometries of each feature set.

In form processing applications where different form layouts are to be processed, form identification (classification) can be addressed using the same idea presented in this thesis. The problem is now extended to N (number of possible form layouts to be processed) model form and one image form to be processed each time.

Bibliography

1. A. Goshtasby and G. C. Stockman, "Point pattern matching using convex hull edges", *IEEE Trans. Syst. Man Cybern.*, vol SMC-15, pp. 631-637, 1985.
2. P. M. Griffin and C. Alexopolous, "Point pattern matching using centroid bounding", *IEEE Trans. Syst. Man Cybern.*, vol. 19, No. 5, pp. 1274-1276, 1989.
3. J. Sprinzak and M. Werman, "Affine point matching", *Pattern recognition letter*, pp. 337-339, April 1994.
4. V. V. Vinod and S. Ghose, "Point matching using asymmetric neural networks", *Pattern recognition*, vol. 26, no. 8, pp. 1207-1214, 1993.
5. Hu, "Visual pattern recognition by moment invariants", *Proc. IRE-49*, pp. 113-121, 1962.
6. I. Herman , "The Use Of Projective Geometry In Computer Graphics", *Lecture Notes In Computer Science 564*, Springer-Verlag 1992.
7. K. Wang, "Affine-invariant moment method of three dimensional object identification", *Ph.D. Thesis, Syracuse Univ.*, Syracuse 1977.
8. J. Mundy, A. Zissermann, "Geometric invariance in computer vision", *MIT press*, UK 1992.
9. David Gans, *Transformations and Geometries*. Appleton-Century-Crofts, New York, 1969.

10. W. Postl, "Detection of linear oblique structures and skew scan in digitized documents," in Proc. 8'th Int. Conf. Patt. Recogn. (ICPR) (Paris), pp. 687 - 689, Oct. 1986 .
11. H. S. Baird, "The skew angle of printed documents," in Proc. Conf. Soc. Photog. Scien. Eng. (Rochester , NY), pp. 14 - 21, May 1987.
12. T. Akiyama and N. Hagita , "Automated entry system for printed documents," Patt. Recogn., vol. 23, no. 11, pp. 1141 - 1154, 1990.
13. S.N. Srihari and V. Govindarju, " Machine vision applications," vol. 2, pp. 141 - 153, 1989 .
14. S. C. Hinds, J. L. Fisher, and D. P. D'Amoto, "A document skew detection method using run-length coding and Hough transform," in Proc. 10th Int. Conf. Patt. Recogn. (ICPR), (Atlantic City , NJ), pp. 464 - 468, June 1990.
15. R. Safari, N. Narasimhamurthi, M. Shridhar and M. Ahmadi. "Document registration using projective geometry", Int. Conf. on document analysis and recognition, pp 1164-1167, Aug. 1995.
16. Irene Rothe et al, "A general method to determine invariants", Research Report, Univ. of Rochester, 1994.
17. D.H.Ballard, "Generalizing the Hough Transform to Detect Arbitrary Shapes", Pattern Recognition, 13(2): pp 111-122, 1981.
18. G. Stockman, "Object recognition and Localization via Pose Clustering", J. of Computer Vision, Graphics, and Image Processing, 40(3), pp 361-387, 1987.

19. D.P. Huttenlocher and S.Ullman, "Object Recognition using Alignment", Proc. of ICCV, pp 102 - 111, London, 1987.
20. Y. Lamdan, H. J. Wolfson, "Geometric Hashing : A General and Efficient Model-Based Recognition scheme", Proc. of ICCV, pp 238 - 249, Tampa, Florida, Dec. 1988.
21. H. J. Wolfson, "Model Based Object Recognition by Geometric Hashing", Proc. of ECCV, pp 526 - 536, Antibes, France, April 1990.
22. P. J. Besl and R. C. Jain, "Three Dimentional Object Recognition", ACM Computing Surveys, 17(1), pp 75-154, 1985.
23. R. T. Chin and C. R. Dyer, "Model Based Recognition In Robot Vision", ACM Computing Surveys, 18(1), pp 67-108, 1986.
24. Y. Lamdan, J. C. Schwartz and H. J. Wolfson, "Object Recognition By Affine Invariant Matching", Proc. of CVPR, pp 335-344, Ann Arbor, Michigan, June 1988.
25. Y. Lamdan, J. C. Schwartz and H. J. Wolfson, "On Recognition Of 3-D Objects from 2-D Images", Proc. of IEEE Int. Conf. on Robotics and Automation, pp 1407-1413, Philadelphia, Pa., Apr. 1988.
26. Algorithms in C Sedgewick 1993.
27. R. O. Duda and P. E. Hart, " Use of the Hough Transform To Detect Lines and Curves in Pictures", Communications of ACM, vol. 15, no. 1, pp. 11-15, 1972
28. J. Illingworth and J. Kittler, "A Servey of the Hough Transform", CVGIP, vol. 44, pp. 87-116, 1988.

29. H. Kalviainen, P. Hirvonen, L. Xu, and E. Oja. "Comparison of Probabilistic and Non-Probabilistic Hough Transform", Proc. of 3rd European Conference on Computer Vision ECCV94, Stockholm, Sweden. pp. 351-360, May 1994.
30. B. K. P. Horn, "Robot Vision", The MIT Press, New York, 1986.
31. J. Serra, "Image Analysis and Mathematical Morphology", Academic Press, New York, 1982.
32. F. M. Wahl, K. Y. Wong, and R. G. Casey. "Block segmentation and text extraction in mixed text/image documents", Computer Graphics and Image Processing, vol. 20, pp. 375-390, 1982.
33. J. Toyoda, Y. Noguchi, and Y. Nishimura. "Study of extracting Japanese newspaper article", in Proc. 6th Int. Conf. Pattern Recognition (Munich, Germany), 1982, pp. 1113-1115.
34. M. Okamoto and A. Miyazawa, "An experimental implementation of document recognition system for papers containing mathematical expressions". in Pre-Proc. 1990 Syntactic & Structural Pattern Recognition (Murray Hill, NJ), 1990, pp. 335-351.
35. H. S. Baird and K. Thompson, "Reading Chess". IEEE Transactions Pattern Anal. Mach. Intell., vol. 12, no. 6, pp. 552-559, 1990.
36. F. Spositi, D. Malerba, and G. Semeraro, "An experimental page layout recognition system for office document automatic classification: an integrated approach for inductive generation", in Proc. 10th Int. Conf. Pattern Recognition (Atlantic City, NJ), 1990, pp. 557-562.

37. J. Higashino, H. Fujisawa, Y. Nakano, and M. Ejiri, "A knowledge-based segmentation method for document understanding", in Proc. 8th Int. Conf. Pattern recognition (Paris, France), 1986, pp. 745-748.
38. Y. Tsuji et al. "Document recognition system with layout structure generator", in Proc. IAPR Workshop on Machine Vision Applications (Tokyo, Japan), 1990, pp. 479-482.
39. Y. Nakano, H. Fujisawa, and O. Kunisaki, "A document understanding system incorporating character recognition", in Proc. 8th Int. Conf. Pattern Recognition (Paris, France), 1986, pp. 801-803.
40. K. Inagaki, T. Kato, T. Hiroshima, and T. Sakai, "MACSYM: A hierarchical parallel image processing system for event-driven pattern understanding for documents", Pattern Recognition, vol. 17, no. 1, pp. 85-108, 1984.
41. I. Masuda, N. Hagita, and T. Akiyama, "Approach to smart document reader system", in Proc. 1985 Comput. Vision Pattern Recognition (San Francisco, CA), 1985, pp. 550-557.
42. H. S. Baird, "Anatomy of a page reader", in Proc. IAPR Workshop on Machine Vision Applications (Tokyo, Japan), 1990, pp. 483-486.
43. S. N. Srihari and G. W. Zack, "Document Image Analysis", in Proc. 8th Int. Conf. Pattern Recognition (Paris, France), 1986, pp. 434-436.
44. J. Yuan, L. Xu, and C. Y. Suen, "Form item extraction by model matching", in Proc. of the First Int. Conf. on Document Analysis and Recognition, pp. 210-218, 1991.

45. R. G. Casey and D. R. Fergusen, "Intelligent Form Processing". IBM System Journal, vol 29, pp. 435-450, 1990.
46. S. Liebowitz Taylor, R. Fritzson, and G. A. Pastor, "registration and region extraction of data from forms", Machine Vision and Applications, vol. 5, pp. 211-212, 1992.
47. D. Wang and S. N. Srihari, "Analysis of form images", in Proc. of the first Int. Conf. on Document Analysis and recognition, pp. 181-191, 1991.
48. Lei Xu, "Form Analysis by Model Matching", Internal Research proposal, Center for Pattern recognition and Machine Intelligence, Concordia university, sept. 28, 1990.
49. P. Kultanen, lei Xu and E. Oja, "Generalized Randomized Hough Transform (GRHT)", Research Report no. 20, dept. of Information Tech., Lappeenranta Univ. of Tech., Finland, 1989.
50. Sh. Tsujimoto and H. Asada, "Major components of a text reading system", Proceedings of the IEEE, vol. 80, no. 7, July 1992.
51. T. Pavlidis, and J. Zhou, " Page segmentation by white streams", in Proc. First Int. Conf. Document Anal. Recogn. (ICDAR), Sept. 1991, pp. 945-953.
52. S. N. Srihari, V. Govindaraju, "Analysis of textual images using Hough transform," Machine Vision Applications, vol 2, pp. 141-153, 1989.
53. A. Hashizume, P. S. Yeh, and A. Rosenfeld, "A method of detecting the orientation of aligned components," Patt. Recogn. Lett., vol. 4, pp. 125-132, 1986.

54. L. O'Gorman," the document spectrum for page layout analysis", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol 15, no. 11, pp. 1162-1173, Nov. 1993.

Appendix

Three more test forms are presented in this appendix including the blank form followed by a filled in form and its corresponding registered form and finally the user entered data extracted from the form.

HANDWRITING SAMPLE FORM

NAME	DATE	CITY	STATE	ZIP
<small>This sample of handwriting is being collected for use in testing computer recognition of hand printed numbers and letters. Please print the following characters in the boxes that appear below.</small>				
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9				
00 550 1000 99417 735190				
27 8701 78317 134666 10				
1509 57970 817837 17 183				
13917 664314 85 169 8171				
147787 75 785 4884 90828				
r u e q d e p h b s . w a y o u g f h i n i y s o n				
X K C I T S C E B U H L J P W D Y R X M O F A Q Z V				
<small>Please print the following text in the box below:</small>				
<small>We, the People of the United States, in order to form a more perfect Union, establish Justice, insure domestic Tranquility, provide for the common Defense, promote the general Welfare, and secure the Blessings of Liberty to ourselves and our posterity, do ordain and establish this CONSTITUTION for the United States of America.</small>				

Figure A-1 : First sample of the blank template.

[illegible]

94

Random May 20th 1968 1011 1011
0123456789 0123456789 0123456789
 09 348 5098 98417 735100
222 820 2607 350666 45
1500 52820 017632 17 46
 13417 6K4310 21 89 6125
242282 21 7015 4586 90848
ru99dphzmswzcmfklxtyvsn
YKGLTSCFRUHLPWDYRXM-F42CN

We the People of the United States, in order to
form a more perfect Union, establish Justice,
insure domestic Tranquility, Provide for the
common Defense, Promote the General
Welfare, and secure the Blessings of
Liberty to ourselves and our Posterity do
hereby establish this Constitution for the
United States of America

Figure A-4 : Extracted information from figure A-3.

Revenue Canada
Customs, Excise and Taxation

Revenu Canada
Accès, Douanes et Impôt

T1-00
1993

Direct Deposit Request Form for Individuals

- Complete this form to have your income tax refund, goods and services tax (GST) credit, or child tax benefit payments deposited directly.
- Attach the completed form to your income tax return or send it to the nearest Revenue Canada income tax office.
- For more information on Direct Deposit, see the back of this form.

Part A - Identification (Please print clearly)

First name and initial		Last name		Social insurance number	
Address		City		Province or territory	
Is this a new address?		Yes ☐ No ☐		Postal code	

Part B - To start Direct Deposit or change information

- To have your income tax refund and GST credit deposited directly, enter the banking information in areas 702, 703 and 704:

Branch number	Institution number	Account number
---------------	--------------------	----------------

Give name(s) of account holder(s), if different from Part A:

- To have your child tax benefit deposited directly, check box 701. OR enter the banking information in areas 705, 706 and 707:

☐ I want my benefit deposited in the same account as above.

Branch number	Institution number	Account number
---------------	--------------------	----------------

Give name(s) of account holder(s), if different from Part A:

Part C - To confirm banking information

Attach a personal cheque with "VOID" written on the front. OR have your financial institution complete Part C.

Name, address and postal code of financial institution (A stamp can be used)

Signature of financial institution official
Telephone no. of financial institution

Year Month Day

Name, address and postal code of financial institution (A stamp can be used)

Signature of financial institution official
Telephone no. of financial institution

Year Month Day

Part D - To stop Direct Deposit

Indicate for which payment(s) you wish to stop Direct Deposit (Check one box only):

1 ☐ All payments

2 ☐ Income tax refund and GST credit

3 ☐ Child tax benefit

Part E - Authorization

I, as the person entitled to receive the above-mentioned payment(s) and in lieu of my receiving a cheque for the same from the Receiver General, hereby authorize the Receiver General for Canada to deposit, until further notice, the payment(s) described above into the account(s) noted herein by means of Direct Deposit.

I also agree that if Revenue Canada, Customs, Excise and Taxation is advised by my financial institution of a change in my financial institution number, branch number or account number, it may redirect payments to me in accordance with the information provided by my financial institution.

Signature Telephone number Date

Figure A-5 : Second sample of the blank template.

Figure A-6 : Second sample of the filled in form with geometric distortion.

Revenue Canada and Taxation

Direct Deposit Request Form for Individuals

For more information on Direct Deposit, see the back of this form.

• Complete this form to have your income tax refund, goods and services tax (GST) credit, or child tax benefit payments deposited directly.

• Attach the completed form to your income tax return or send it to the nearest Revenue Canada office.

• For more information on Direct Deposit, see the back of this form.

Part A - Identification (Please print clearly)

Last name and first name: KAMIN
 Address: SAFARI
 City: 310 DONNELLY ST
 Province or territory: ON
 Postal code: M5S 1A5

Is this a new request? ☒ Yes ☐ No

Part B - To start Direct Deposit or change information

Enter your banking information in areas 702, 703 and 704:

Bank: 3355 Branch: 566 Account number: XXXX

Give branch(es) or account number(s), if different from Part A.

☒ I want my benefit deposited in the same account as above.

On your first tax benefit deposited directly, check box 707.

On your first tax benefit deposited directly, check box 707.

Part C - To confirm banking information

Attach a preauthorized cheque with "VOID" written on the front. On your first request, attach Part C.

Name, address and postal code of financial institution (LA stamp can be used):
 Signature of financial institution official (LA stamp can be used):
 Telephone no. of financial institution (LA stamp can be used):

Part D - To stop Direct Deposit

Indicate for which payment(s) you wish to stop Direct Deposit (Check one box only):

☒ All payments ☐ Income tax refund and GST credit ☐ Child tax benefit

Part E - Authorization

I, as the person entitled to receive the above-mentioned payment(s), and in lieu of my receiving a cheque for the same from the Receiver General, hereby authorize the Receiver General for Canada to deposit, and further receive, the payment(s) described above and the account(s) used herein by means of Direct Deposit.

I also agree that Revenue Canada, Customs, Excise and Taxation is advised by my financial institution of a change in my financial institution number, branch number or account number. I may request payment to me in accordance with the information provided by my financial institution.

Signature: [Signature]
 Date: 9/1/99
 Telephone number: 823473589

Direct Deposit Request Form for Individuals

- Complete this form to have your income tax refund, goods and services tax (GST) credit, or child tax benefit payments deposited directly.
- Attach the completed form to your income tax return or send it to the nearest Revenue Canada income tax office.
- For more information on Direct Deposit, see the back of this form.

Part A - Identification (Please print clearly)	
First name and name Address City Province or territory	Last name Social insurance number Is this a new address? Yes ☒ No ☐ Postal code
RAMIN 4605 3170 DOWNS FILL ST WINDSOR ONTARIO L9A 4K6	SAFARI 1-1 1-1 1-1 Yes ☒ No ☐ L9A 4K6
Part B - To start Direct Deposit or change information	Part C - To confirm banking information
To have your income tax refund and GST credit deposited directly enter the banking information in areas 702, 703 and 704. Branch number Institution number Account number 3355 566161 X Give name(s) of account holder(s), if different from Part A. To have your child tax benefit deposited directly, check box 701. Or enter the banking information in areas 705, 706 and 707. ☒ I want my benefit deposited in the same account as above. Branch number Institution number Account number Give name(s) of account holder(s), if different from Part A.	Attach a personalized cheque with "MIND" written on the front. Or have your financial institution complete Part C. Name, address and postal code of financial institution (A stamp can be used) Signature of financial institution official Telephone no. of financial institution Date Year Month Day 1973 5 4 9 6 13 1 Name, address and postal code of financial institution (A stamp can be used) Signature of financial institution official Telephone no. of financial institution Date Year Month Day
Part D - To stop Direct Deposit	
Indicate for which payment(s) you wish to stop Direct Deposit (Check one box only): 1 ☒ All payments 2 ☒ Income tax refund and GST credit 3 ☒ Child tax benefit	
Part E - Authorization	
I, as the person entitled to receive the above-mentioned payment(s) and in lieu of my receiving a cheque for the same from the Receiver General, hereby authorize the Receiver General for Canada to deposit, until further notice, the payment(s) described above via the account(s) noted herein by means of Direct Deposit. I also agree that if Revenue Canada, Customs, Excise and Taxation is advised by my financial institution of a change in my financial institution number, branch number or account number, I may redirect payments to me in accordance with the information provided by my financial institution. Signature: R. Safari Telephone number: 93347315891 Date: Year Month Day 9 6 1	

Figure A-7 : Second sample of the registered image to its blank template.

RAMIN SAFARI 1-1 1-1 1-1
#608 3170 DOWVELLY ST
WINDSOR ONTARIO 1-1-1-1

3355 5555111

4735191 940731

x

R. Lefebvre 923473589, 94

Figure A-8 : Extracted information from figure A-7.

OFFICE OF THE REGISTRAR
 GRADE COLLECTION FORM

SEMESTER

PAGE

01

EE 525 01 (06 88) 2-D DIGITAL SIGNAL PROCE

TBA

TBA

	STUDENT	STUDENT	COMMENTS	GRADE	DEGREE
1	AMHED, ZULFIJAR	969-010275		M2	EE MAS
2	BEZAIRE, ANDY P.	921-144660		M2	EE MAS
3	CONFLITTI, DANNY	921-204450		M2	EE MAS
4	FRANKLIN, TRACY A.	921-133730		M2	EE MAS
5	JOVANOVIC, MARKS	800-300980		M1	EE MAS
6	LEE, ALFRED	920-784100		M2	EE MAS
7	MIAN, TARIQ M.	969-010205		M2	EE MAS
8	MUSCEDERS, ROBERT J	921-135800		M2	EE MAS
9	RAAHENIFAR, RAAMRAH	969-010244		02	EE PHD
TOTAL STUDENTS = 9					

GRADES MUST BE SUBMITTED TO DEAN'S
 OFFICE NO LATER THAN JAN 3, 1997.

2454

PROFESSOR'S SIGNATURE

DEPARTMENT HEAD'S SIGNATURE

DEAN'S SIGNATURE

Figure A-9 : Third sample of the blank template.

OFFICE OF THE REGISTRAR
GRADE COLLECTION FORM

1998 FALL
University of Windsor
SEMESTER

96/11/28

EE 325 01 (06 88) 2-D DIGITAL SIGNAL PROC

TBA

COMMENTS

PAGE 01

1 AMIR, ZULFIJAH
2 BEZURE, ANDY P.
3 COMPELLITI, DANNY
4 FRANKLIN, TRACY A.
5 JOVACHOVIC, MARKO
6 LEE, ALFRED
7 MIANO, TARIQ A.
8 MUJCEDEBAC, ROBERT J
9 NAHMEHIFAK, KAMRAN

TOTAL STUDENTS = 9

969-010275 A	M2	EE	MAS
921-144660 B	M2	EE	MAS
921-264450 C	M2	EE	MAS
921-133730 D	M1	EE	MAS
800-300980 A	M2	EE	MAS
920-784100 C	M2	EE	MAS
969-010205 B	M2	EE	MAS
921-135800 D	M2	EE	MAS
969-010244 A	O2	EE	PHD

DEPARTMENT HEAD'S SIGNATURE

2454 G. H.

PROFESSOR'S SIGNATURE

H. R.

DEAN'S SIGNATURE

0.12

GRADES MUST BE SUBMITTED TO DEAN'S OFFICE NO LATER THAN JAN 3, 1997

Figure A-10 : Third sample of the filled in form with geometric distortion.

OFFICE OF THE REGISTRAR
GRADE COLLECTION FORM

PAGE

01

EE 525 01 (06 88) 2-D DIGITAL SIGNAL PROCE

TBA

TBA

STUDENT	STUDENT ID	EXAMINATIONS	GRADE	MARKS	REMARKS
1 AHMED, ZULFIKAR	969-010275 A		M2	EE	MAS
2 BEZAIRE, ANDY P.	921-144660 B ⁺		M2	EE	MAS
3 CONFLITTI, DANNY	921-204450 C ⁻		M2	EE	MAS
4 FRANKLIN, TRACY A.	921-133730 D ⁻		M2	EE	MAS
5 JOVANDVIC, MARKO	800-300980 A ⁻		M1	EE	MAS
6 LEE, ALFRED	920-784100 C ⁺		M2	EE	MAS
7 MIAN, TARIQ A.	969-010205 B ⁻		M2	EE	MAS
8 MUJCEDELC, ROBERT J	921-135800 D ⁻		M2	EE	MAS
9 RAHMANIFAR, RAHMAN	969-010244 A		O2	EE	PHD

TOTAL STUDENTS = 9

GRADES MUST BE SUBMITTED TO DEAN'S
OFFICE NO LATER THAN JAN 3, 1997.

 PROFESSOR'S SIGNATURE

2454 G. K.
 DEPARTMENT HEAD'S SIGNATURE

 DEAN'S SIGNATURE

Figure A-11 : Third sample of the registered image to its blank template.

A			
B'			
C'			
D'			
A'			
C'			
B			
D'			
A	K. Rastu	G. Rastu	D. Rastu

Figure A-12 : Extracted information from figure A-11.

PUBLICATIONS

Journals

- [1] R. Safari, N. Narasimhamurthi, M. Shridhar and M. Ahmadi, "Document Registration Using Projective Geometry", IEEE Transactions on Image Processing, Vol. 6, pp. 1337-1341, Sept. 1997.

Workshops

- [1] R. Safari, H. Djahanshahi, S. Sadeghi, M. Ahmadi, G. A. Jullien, and W. C. Miller, "ANN's and CNN's in New Applications for Pattern Recognition and Image Processing", Micronet Annual Workshop, Ottawa, Ontario, Canada, pp 71-72, Mar. 1997.

Conferences

- [1] R. Safari, N. Narasimhamurthi, M. Shridhar and M. Ahmadi, "Document Registration Using Projective Geometry", Int. Conf. on Document Analysis and Recognition, pp 1164-1167, Aug. 1995.
- [2] R. Safari, N. Narasimhamurthi, M. Shridhar and M. Ahmadi, "Extraction of Distortion Parameters in Forms and Documents", Second Asian Conference on Computer Vision, Singapore, Vol. III, pp 131-135, Dec. 1995.
- [3] R. Safari, N. Narasimhamurthi, M. Shridhar and M. Ahmadi, "Affine Point pattern Matching and its Application to Form Registration", IEEE Int. Symposium on Circuit and Systems, Atlanta, USA, Vol. II, pp 656-659, May. 1996.
- [4] R. Safari, N. Narasimhamurthi, M. Shridhar and M. Ahmadi, "Form Registration : A Computer Vision Approach", Int. Conf. on Document Analysis and Recognition, pp 758-761, Aug. 1997.

VITAE AUCTORIS

- 1964 Born in Sept, Tehran, Iran.
- 1987 Received B. A. Sc. degree in Electronics from Shiraz University, Shiraz, Iran.
- 1989 Received M. A. Sc. degree in Digital Communication from Isfahan University of Technology, Isfahan, Iran.
- 1993 Candidate for Ph. D. degree in Electrical Engineering at University of Windsor, Windsor, Ontario, Canada.