

Rochester Institute of Technology

## RIT Digital Institutional Repository

---

### Theses

---

1987

## Adaptive statistical recognition of hand-printed Telugu characters

Murthy Lakshmana Mantha

Follow this and additional works at: <https://repository.rit.edu/theses>

---

### Recommended Citation

Mantha, Murthy Lakshmana, "Adaptive statistical recognition of hand-printed Telugu characters" (1987). Thesis. Rochester Institute of Technology. Accessed from

This Thesis is brought to you for free and open access by the RIT Libraries. For more information, please contact [repository@rit.edu](mailto:repository@rit.edu).

ROCHESTER INSTITUTE OF TECHNOLOGY  
SCHOOL OF COMPUTER SCIENCE AND TECHNOLOGY

ADAPTIVE STATISTICAL RECOGNITION OF  
HAND-PRINTED TELUGU CHARACTERS

BY

MURTHY LAKSHMANA MANTHA

This thesis is submitted to The Faculty of the School of  
Computer Science and Technology in partial fulfillment of  
the requirements for the degree of Master of Science in  
Computer Science.

Approvals:

**John A. Biles**

-----  
Thesis Advisor: Professor John A. Biles

**L. A. Coon**

-----  
Committee Member: Professor Larry Coon

**Peter G. Anderson**

-----  
Committee Member: Professor Peter Anderson

## ABSTRACT

A brief description of statistical and syntactic pattern matching techniques is presented with an emphasis on statistical techniques. The characteristics of the Telugu script are described. A subset of 16 characters, which are both easy and hard to recognize, is selected for the dictionary of standard characters. A weighted linear difference polynomial of features is used to recognize Telugu characters. The features were Fourier descriptors of projection profiles and cross sections taken in various directions. Algorithms for obtaining the projection profiles, cross sections and adaptive learning method are presented. The system was trained and tested with a set of 8 nano-written samples of each of 16 different Telugu characters. More than 90% of the 128 samples were correctly recognized by the system. Results of numerous trials examining the different features and classification techniques are discussed.

## ACKNOWLEDGEMENTS

I wish to acknowledge my thanks to Jeff Peltz, Faculty Member, Imaging Science Department, for making the digital video equipment available for this study; Dave Medvedeff, at the ROSS Computing Center for helping me with VAX/VMS system, and Al Biles, my thesis advisor, for his excellent guidance and critical review of the manuscript. Last but not least, I would like to thank my friend Dersy Song, whose thesis work in Chinese character recognition has been a source of inspiration and reference to this study.

## TABLE OF CONTENTS

- 1. INTRODUCTION AND OVERVIEW
  - 1.1 INTRODUCTION TO PATTERN RECOGNITION
  - 1.2 ABOUT TELUGU CHARACTERS
  - 1.3 OBJECTIVE
- 2. THEORITICAL AND CONCEPTUAL DEVELOPMENT
  - 2.1 TEMPLATE MATCHING
  - 2.2 STATISTICAL METHODS IN PATTERN RECOGNITION
    - 2.2.1 FEATURE EXTRACTION
      - 2.2.1.1 METHOD OF MOMENTS
      - 2.2.1.2 PROJECTION PROFILES
      - 2.2.1.3 CROSS-SECTIONS
    - 2.2.2 CLASSIFICATION METHODS
    - 2.2.3 LEARNING
  - 2.3 SYNTACTIC METHODS IN PATTERN RECOGNITION
    - 2.3.1 LINGUISTIC PATTERN RECOGNITION SYSTEM
    - 2.3.2 SELECTION OF PATTERN PRIMITIVES
    - 2.3.3 PATTERN GRAMMARS
    - 2.3.4 SYNTAX ANALYSIS AS RECOGNITION PROCEDURE
  - 2.4 DESCRIPTIVE CHARACTERISTICS OF TELUGU CHARACTERS
    - 2.4.2 PREVIOUS WORK IN TELUGU CHARACTER RECOGNITION
    - 2.4.1 RECOGNITION OF TELUGU CHARACTERS USING  
STRUCTURAL METHODS
- 3. RELEVANT THEORIES AND CONCEPTS
  - 3.2 FEATURE SELECTION AND EXTRACTION
    - 3.2.1 PROJECTION PROFILES

3.2.2	CROSS-SECTIONS OR RUNS
3.2.3	CONDENSED CROSS-SECTIONS
3.2.4	NOISE REMOVAL IN CONDENSED STRINGS
3.3	MAGNITUDE SPECTRA USING FOURIER ANALYSIS
3.4	ADAPTIVE LEARNING
3.5	CLASSIFICATION
4.	IMPLEMENTATION AND EXPERIMENTAL PROCEDURES
4.1	IMPLEMENTATION
4.2	HARDWARE CONFIGURATION
4.3	SOFTWARE CONFIGURATION
4.4	TEST DATA
4.5	TEST PROCEDURES
5.	RESULTS AND DISCUSSION
5.1	RESULTS OF TEST 1
5.2	RESULTS OF TEST 2
5.3	RESULTS OF TEST 3
5.4	RESULTS OF TEST 4
5.5	RESULTS OF TEST 5
6.	CONCLUSION AND FUTURE RESEARCH
6.1	CONCLUSION
6.2	FURTHER STUDY
7.	BIBLIOGRAPHY
8.	ACKNOWLEDGEMENTS
9.	LIST OF ILLUSTRATIONS

## LIST OF ILLUSTRATIONS

- 1.1 SAMPLE TELUGU PRINT
- 2.1 STATISTICAL PATTERN RECOGNITION SYSTEM
- 2.2 SYNTACTIC PATTERN RECOGNITION SYSTEM
- 2.3 CROSS-SECTIONS FOR A TAMIL CHARACTER
- 2.4 CROSS-SECTIONS FOR NUMERALS 0-9
- 2.5 BLOCK DIAGRAM OF A PATTERN CLASSIFIER
- 2.6 PATTERN CLASSIFICATION BY PROXIMITY CONCEPT
- 2.7 PRIMITIVES FOR A RECTANGLE
- 2.8 RECTANGLE WITH PRIMITIVES IN FIG. 2.7
- 2.9 TREE STRUCTURE FOR 2.7
- 2.10 STRUCTURAL DESCRIPTION OF A PATTERN FROM FIG.2.7
- 2.11 DIRECTED LABELLED GRAPH FOR TAMIL CHARACTER
- 2.12 BLOCK DIAGRAM OF LINGUISTIC PATTERN RECOGNITION SYSTEM
- 2.13 BUILD PRIMITIVES FOR DEVANAGARI SCRIPT
- 2.14 COMPOSITION OF 'क' AND ITS PLANG DESCRIPTION
- 2.15 LABELLED PATTERN AND PARSE TREE FOR 'क'
- 2.16 FREEMAN CHAIN CODE
- 2.17 PATTERN PRIMITIVES IN TERMS OF REGIONS
- 2.18 CONTEXT SENSITIVE AND CONTEXT FREE GRAMMARS FOR A RECTANGLE
- 2.19 REPRESENTATION OF CHROMOSOMES
- 2.20 REPRESENTATION OF A HOUSE
- 2.21 HEAD AND TAIL PRIMITIVES

2.22	PICTURE DEFINITION LANGUAGE FOR A 'HOUSE'
2.23A	TELUGU VOWELS
2.23B	CONSONANTS
2.23C	VOWEL SIGNS
2.23D	EXAMPLES OF C-V LETTERS
2.23E	EXAMPLES OF C-V LETTERS
2.23F	BASIC LETTERS
2.23G	BUILD PRIMITIVES
2.23H	CONJUNCT PRIMITIVES
2.23I	CONJUNCT CONSONANTS WITH CONJUNCT PRIMITIVES
2.23J	CONJUNCT CONSONANTS IN WHICH CONSONANTS WITHOUT THE FIRST VOWEL SIGN APPEAR BELOW C-V LETTERS.
2.24	BLOCK DIAGRAM OF SYNTACTIC RECOGNITION SYSTEM FOR TELUGU CHARACTERS
2.25	SAMPLE T-TUPLES
2.26	TELUGU BUILD PRIMITIVE AND ITS T-TUPLES
2.27	CODE IN TERMS OF T-TUPLES FOR TELUGU '✓'
3.1	FUNCTIONAL BLOCK DIAGRAM
3.2	ALGORITHMS FOR OBTAINING PROJECTION PROFILES
3.3	PROJECTION PROFILES OF CHARACTER 'అ'
3.4	CROSS SECTION OF CHARACTER 'అ'
4.1	HARDWARE CONFIGURATION
4.2	SOFTWARE CONFIGURATION
4.3	HAND-PRINTED STANDARD CHARACTERS
4.4	HAND-PRINTED SAMPLE CHARACTERS 1,2
4.5	HAND-PRINTED SAMPLE CHARACTERS 3,4
4.6	HAND-PRINTED SAMPLE CHARACTERS 5,6
4.7	HAND-PRINTED SAMPLE CHARACTERS 7,8



- 4.8 HAND-PRINTED SAMPLE CHARACTERS 9,10
- 4.9 HAND-PRINTED SAMPLE CHARACTERS 11,12
- 4.10 HAND-PRINTED SAMPLE CHARACTERS 13,14
- 4.11 HAND-PRINTED SAMPLE CHARACTERS 15,16
- 4.12 DATA FILE STRUCTURES FOR STANDARD CHARACTERS
- 4.13 DATA FILE STRUCTURES FOR SAMPLE CHARACTERS
- 5.1 SUMMARY OF TEST RESULTS
- 5.2 CROSS-SECTION VALUES AND FOURIER DESCRIPTORS OF  
PROJECTION PROFILES FOR 16 STANDARD CHARACTERS
- 5.3 FOURIER DESCRIPTORS OF PROJECTION PROFILES AND  
CROSS-SECTIONS FOR 16 STANDARD CHARACTERS
- 5.4 CONFUSION MATRIX FOR RESULTS OF TEST-1
- 5.5 CONFUSION MATRIX FOR RESULTS OF TEST-2
- 5.6 CONFUSION MATRIX FOR RESULTS OF TEST-3
- 5.7 CONFUSION MATRIX (COLUMNS) FOR RESULTS OF TEST-4  
AND 5
- 5.8 CONFUSION MATRIX (ROWS) FOR RESULTS OF TEST-4  
AND 5
- 5.9 CONDENSED CROSS-SECTIONS FOR STANDARD AND SAMPLE  
CHARACTERS
- 5.10 CONDENSED CROSS-SECTIONS AFTER NOISE REMOVAL FOR  
STANDARD AND SAMPLE CHARACTERS

## CHAPTER 1

### INTRODUCTION AND OVERVIEW

#### 1.1 Introduction To Pattern Recognition.

The problem of pattern recognition has received the attention of many researchers for the past 25 years. The approaches taken to solve the problem of pattern recognition fall into three general categories: (1) template matching, (2) the classificatory approach and (3) the syntactic or structural approach.

In the template matching method, an unknown pattern is compared with the template of each pattern classes and the classification is based on preselected matching criteria.

In the classificatory approach a set of characteristic measurements, called 'features', is extracted from the patterns, and the feature set is treated as a point in a multidimensional space. The recognition of each pattern is usually made by partitioning the feature space into subspaces, each subspace corresponding to a pattern class. This partitioning, made by decision functions, might be effected on the basis of statistical or non-statistical considerations. This approach has been successfully used on problems like character recognition, medical diagnosis, crop classification etc. The most important aspect of this model

is the a priori selection of an optimum feature set and decision functions. This can be a difficult task even for a simple pattern set. For large and complex sets like Chinese ideographs and Telugu characters, it is not apparent how this selection should be made.

In the syntactic approach, the central aim is to generate the description of a given input in terms of its subparts called 'primitives'. Pattern primitives are selected a priori and their relations in the patterns are described by a grammar. The recognition process is accomplished by parsing the sentences describing the given input patterns. Some of the problems associated with this approach are: the selection of primitives is similar in nature to feature selection; the selection of a suitable shape analyzer to recognize primitives is no easy task; and the selection of a suitable grammar to describe the patterns is difficult.

The key to pattern recognition does not lie wholly in statistical approaches, heuristic programming, or more formal linguistic approaches alone. A good pattern recognition system uses statistical, structural and heuristic tools at various stages of processing of the patterns, with each tool being applied at the stage to which it seems best suited [Kan 72].

## 1.2 About Telugu Characters.

Telugu is one of the national languages of India, spoken by about 60 million people in the state of Andhra Pradesh. The language has a beautiful cursive script with well over 2000 characters. Telugu is a highly phonetic language with each character representing a syllable.

The Telugu alphabet consists of 16 vowels and 36 consonants, consonant-vowel combinations (C-V letters) and conjunct consonants. A vowel following a consonant takes on a different graphic form called a 'vowel sign'. Hence, the vowels and consonants combine together to form 576 different C-V letters. The C-V letters are formed by adding vowel signs to consonants at appropriate places. The C-V letters and consonants are combined in many ways to produce thousands of different characters. Figure 1.1 shows an extract from a Telugu book which illustrates some of the commonly used styles and fonts.

The Telugu characters are not of uniform dimensions. However, they can be conceived as different combinations of special symbols, called primitives. The telugu typewriter keyboard was designed under similar assumptions.

### 1.3 Objective.

The objective of this thesis was to investigate classificatory methods for recognizing hand printed telugu characters.

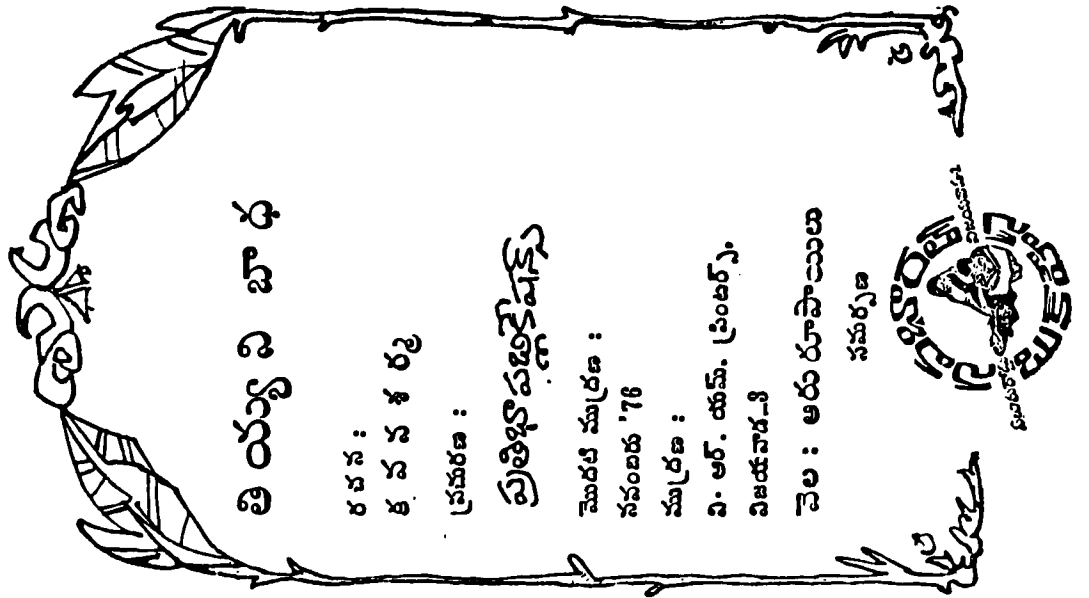


FIG 1.1 SAMPLE TELUGU PRINT

నవలా సాహిత్యంలో కొత్త మార్పు కావా  
లని పాఠకులు కోరినప్పుడలా మేము వారి  
అభిరుచులకు అనుగుణమైన రచనలను  
అందించడంలో అగ్రగణ్యులౌ వున్నాం.  
వతనాన కి అపారంగా పెంచే కృషిలో మా  
వంతు దాద్యత మేము నదా నిర్వహిస్తున్నాం.

ఏలా పెరుగుతున్న పాఠకుల సంఖ్యకు ఏలా వస్తున్న  
సుప్రసిద్ధ రచయితల రచనల సంఖ్య దాదాపులేదని మొదటి  
సారిగా మేమే గురించేము. కనుక, మేము ప్రచురించిన  
ప్రత్యాక రచనలకు దేటుగా వుండే వైవిధ్యంగల ఉత్తమ రచన  
లను వర్తమాన రచయితల నుంచి ఎన్నికచేసి ఒక సాహితీ ఉద్య  
మంగా పాఠకులకు అందిస్తున్నాం.

ఈ ఉద్యమం క్షయమన నవలాసాహిత్యంలో దానంపిన పెద  
మార్పుకి ఒక చిన్న ప్రయత్నం మాత్రమే! ఈ ప్రయత్న  
శాసనీక పాఠకులు నిండుమనసుతో స్వాగతం వయకుతారని  
ఆశిస్తున్నాం. ఆ ప్రయత్నంలో దాగమే—

యూనివర్సిటీ విద్యార్థి విద్యార్థినుల  
ప్రణయ సాహస గాథాలవారి  
కవనశర్మ కావ్య నవలా కథనం

తియ్యవి బాధ

Two different methods of extracting features were used, namely: fourier descriptors of vertical, horizontal, left and right diagonal projection profiles, and fourier descriptors of row and column cross-sections.

Adaptive learning methods were used to train the system with different samples to adjust the weights of features according to recognition or misrecognition.

Very encouraging results ( more than 90% recognition rate) were obtained by using features extracted from a combination of projection profiles in several directions and cross-sections in vertical and horizontal directions. Adaptive learning method improved the performance of recognition by about 20%.

Chapter 2 provides the general background in various techniques of pattern recognition and chapter 3 describes in detail the relevant theories and algorithms used in this study. Chapter 4 describes the implementation methods and test procedures. Chapter 5 contains a discussion of the results. Conclusions and suggestions for further research are discussed in chapter 6.

## CHAPTER 2

### THEORETICAL AND CONCEPTUAL BACKGROUND

The techniques used to solve pattern recognition problems can be grouped into three general approaches; namely, template matching, the statistical (or decision theoretic) approach and the syntactic (or structural) approach. Figure 2.1 shows the typical statistical pattern recognition system which consists of a feature extractor and pattern classifier using the feature measurements from the input pattern. The syntactic pattern recognition system, as shown in Figure 2.2, consists of pattern preprocessing, primitive selection and syntax analysis. This chapter provides an overview of these methods in pattern recognition.

The first step in any pattern recognition problem is to select discriminatory features representing the pattern and to extract (measure) these features. For example, the most important features of handwritten characters are the direction of the strokes, the arrangement of strokes and the interrelation between the strokes. These features, in general, may not be easily measurable. Usually, a binary image of the pattern can be easily obtained and then is preprocessed to extract significant features.

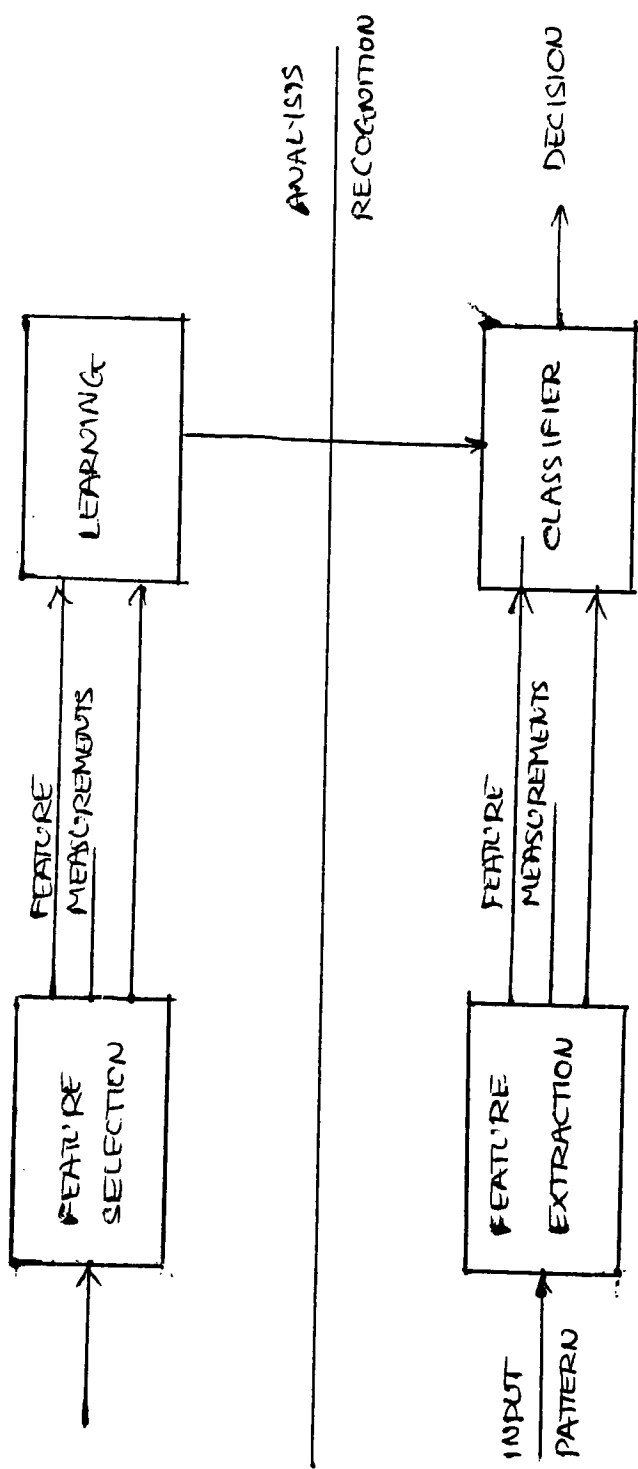


FIG. 2.1 STATISTICAL PATTERN RECOGNITION SYSTEM



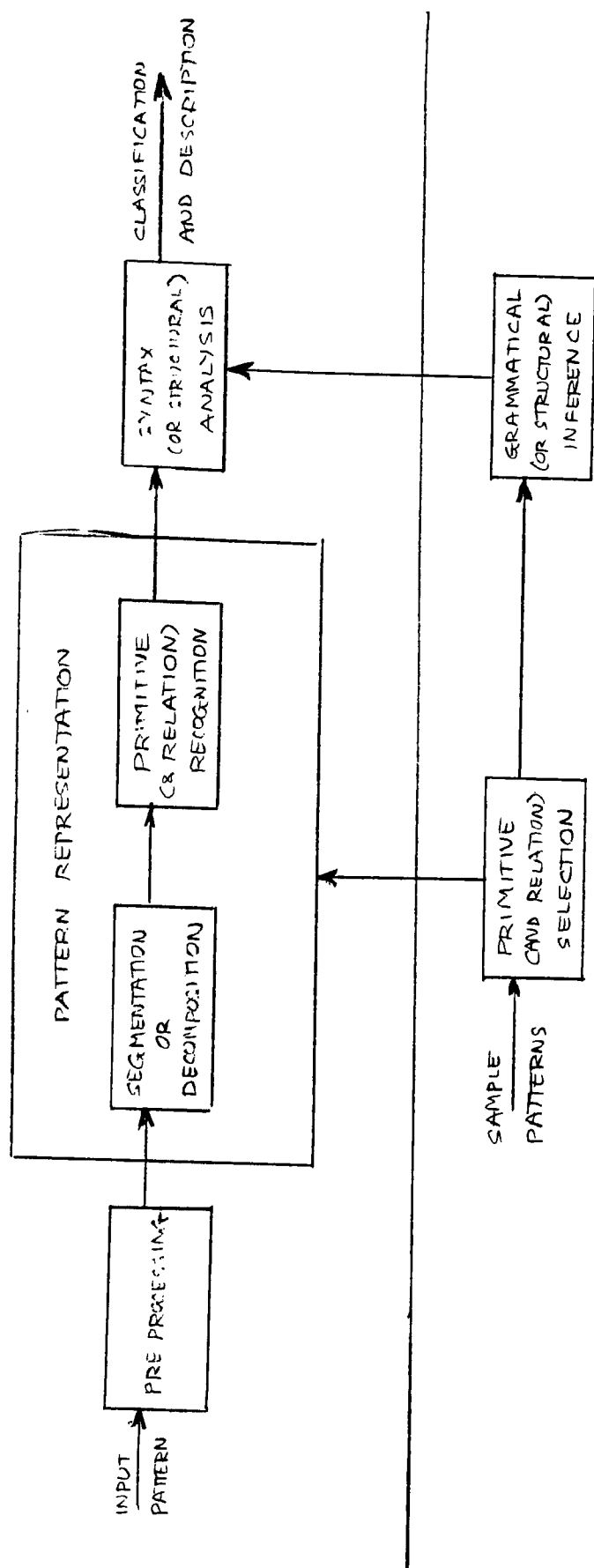


FIG 2.2 BLOCK DIAGRAM OF A SYNTACTIC PATTERN RECOGNITION SYSTEM.

The features can be classified into three categories: (1) physical features, e.g., color (2) structural features, e.g., shape, structure and other geometrical properties and (3) mathematical features, (e.g., statistical means, correlation coefficients, eigen values and eigen vectors of covariance matrices, and other invariant properties). It is generally difficult for a machine to imitate the sensory capability of physical features. Syntactic or structural pattern recognition methods have been developed to analyze structural features. The decision theoretic or classificatory approach is generally used for patterns where features are extracted using mathematical methods.

## 2.1 Template Matching

An intuitive approach to pattern recognition is "template-matching" [Ros 76], [Ben 84]. In this case, a set of templates or prototypes, one for each pattern class, is selected. An unknown input pattern is compared with the template of each classes and the classification is based on preselected matching or similarity criteria. In other words, if the input pattern matches the template of the  $i$ th pattern class better than it matches any other templates, then the input is classified as a member of the  $i$ th pattern class. This approach has been used for some existing printed-character recognizers and bank-check readers.

The following are standard measures of the degree of match between a picture  $f(x,y)$  and a template  $T$  whose gray level at  $(x,y)$  is  $t(x,y)$ :

$$\max \text{ of } |f(x,y) - t(x,y)|, \text{ } x \text{ and } y \text{ in } T \quad (2.1)$$

$$\iint_T |f(x,y) - t(x,y)| \, dx \, dy \quad (2.2)$$

$$\iint_T [f(x,y) - t(x,y)]^2 \, dx \, dy \quad (2.3)$$

These measures are all zero for a perfect match, and have high values for poor matches. They differ, however, in the types of errors to which they are sensitive. For example, if  $f(x,y) = t(x,y)$  except at few points, where  $|f(x,y) - t(x,y)|$  is large, measure 2.1 will show a large mismatch, whereas measures 2.2 and 2.3 will show negligible mismatches. On the other hand, if  $|f(x,y) - t(x,y)|$  is everywhere non-zero but small, e.g.,  $f(x,y) = t(x,y) + e$ , then measure 2.1 yields a small mismatch,  $e$ , while measures 2.2 and 2.3 yield large mismatches  $e|T|$  and  $e^2|T|$ , respectively where  $|T|$  is the area of the template  $T$ .

Another form of measure 2.4 is the correlation coefficient :

$$\frac{\iint_T f(x,y) t(x,y) \, dx \, dy}{\sqrt{\iint_T f^2(x,y) \, dx \, dy \iint_T t^2(x,y) \, dx \, dy}} \quad (2.4)$$

The value of this function is always between 0 and 1, with value 1 being achieved only if  $f(x,y) = ct(x,y)$  where  $c$  is a positive constant.

The disadvantages of the template matching approach include the difficulty in selecting a good template for each pattern class and in defining a proper matching criterion. The difficulty is especially remarkable when large variations and distortions are expected in all the patterns belonging to one class [wid 74]. The use of flexible template matching or "rubber-mask" technique has been proposed by [wid 74] and [Che 75], in an attempt to accommodate large variations in pattern samples. Also, the process of matching a template against a picture in all possible positions is computationally costly. Some methods of reducing the cost of template matching can be found in [Bar 72] and [Mag 72].

## 2.2 Statistical Methods In Pattern Recognition

In this approach, a set of characteristic measurements are extracted from the original pattern and classification is done by partitioning the measurement space into regions of pattern classes, using decision functions derived from pattern samples. Clustering transformations can be used on the measurement space in order to cluster the points representing the same pattern class. Such a transformation will maximize the mean-square distance between pattern points that belong to two different classes and minimize the mean-square distance between pattern points of the same class.

Alternatively, properties of patterns like moments, projections and cross-sections, can be used as characteristic features, and suitable similarity functions can be chosen for classification.

### 2.2.1 Feature Extraction -

The basic idea of feature extraction is to obtain unique information representing the pattern from the gray level in the pattern at different points.

#### 2.2.1.1 Method Of Moments -

In the method of moments used by Dudani, et al [Dud 77] for recognizing aircraft, the picture under study is divided into cells, and the second moment is computed for each cell. The second moment (variance) is defined as the sum of the products of vertical and horizontal positions of gray levels in a given cell. A pattern then can be represented by moments. Moments will be the same for identical patterns and different for dissimilar patterns. By observing the values of the moments, it can be determined whether the gray values are spread out horizontally or vertically, however this method is inadequate for describing structural information of complex patterns [Luc 93].

### 2.2.1.2 Projection Profiles -

A good idea of how the gray levels in a given region of a pattern are distributed can be obtained by examining the projections of the patterns in various directions. The projections in x and y directions are given by:

$$X(i) = \sum_{j=1}^N P(i,j) \quad \text{and} \quad Y(i) = \sum_{j=1}^N P(j,i).$$

A set of projections in a sufficient number of directions contains enough information to reconstruct the picture. For objects having higher gray levels than their backgrounds, peaks in projections can indicate locations of major parts of objects. This is especially true when the image is normalized to contain only 0's and 1's. This method, then, should be well suited for character recognition. Song [Son 85] and Li [Li 83], used Fourier descriptors of projection profiles in horizontal and vertical directions as features for Chinese character recognition. A very high success rate was obtained due to the fact that the characters were composed mainly of vertical and horizontal strokes, and the projections represented these components very well. In addition, the positional invariance of the characters was achieved by deriving the Fourier descriptors, and stroke width correction was applied using the convolution technique of the Fourier transform.

In this thesis the projection profiles of Telugu characters in four directions namely, horizontal, vertical, -45 and +45 degrees were used as features. A combination of the four sets of features yielded unique characteristics.

Details of this method are described in chapter 3.

### 2.2.1.3 Cross-sections -

More detailed information about the arrangement of gray levels in a region can be obtained by examining the cross-sections in various directions (e.g., the cross-section in the x-direction is the function  $f(x, y')$  for a particular value of  $y'$ ). For a binary image matrix with only 0's and 1's, the cross-sections will simply be the number of distinct sequences of runs of 1's. Figure 2.4 shows an example of cross-section values in horizontal direction for numerals 0 through 9. Figure 2.5 illustrates the cross-section values of a Tamil character. Peaks in the cross-sections correspond to object parts. Comparison of cross-sections of different pictures can give useful information about object shape in terms of how the peaks shift, expand, shrink, merge and split.

Dutta [Dut 74] used the change of runs of 1's in horizontal and vertical directions to derive 'events', 'half events' etc. in the recognition of handwritten characters. Siromoney [Sir 78] used the frequency of runs of 1's in columns and rows for recognition of printed Tamil characters. The method is described below:

The pattern matrix is examined column by column and the number of runs of 1's is noted for each column. This gives a string of numbers. This string is condensed by deleting





the consecutive occurrences of the same numbers. This condensing procedure shortens the string considerably and is equivalent to a sort of thinning procedure. Similarly, the matrix is examined row by row and another condensed string is formed. The condensed strings will be the same for identical patterns.

The same method was enhanced to incorporate the relative lengths of the pattern segments. As in case of earlier method, the number of 1's is noted. In this string, any one numeral may occur in consecutive positions. Short, medium and long runs are formed depending on the number of consecutive positions of the same numeral. For example, if the length of the pattern matrix is, say, 12, then 1, 11, 111, will form a short run; 1111, 11111, 111111 will form a medium run and the rest will form a long run. This was called the symbolic run method and gave unique representations for Tamil characters [Sir 78].

The symbolic run method, however, is very sensitive to individual differences of handwritten characters since small variations in writing can cause large variations in cross-sections. However, this method seems to work very well for patterns with horizontal and vertical strokes and with certain connectivity properties. For this thesis, the condensed run method was examined in detail for the recognition of telugu characters, and the details are described in Chapter 3.

### 2.2.2 Classification Methods -

Mathematically, the problem of pattern classification can be formulated in terms of "discriminant" or "decision" functions. Let  $w_1, w_2, \dots, w_m$  be designated as ' $m$ ' pattern classes to be recognized. The pattern space, then, can be considered as consisting of ' $m$ ' regions, each of which encloses the pattern points in a class. The recognition problem can now be viewed as that of generating the decision functions,  $d_1(X), d_2(X), \dots, d_m(X)$ . These functions are scalar and single-valued functions of pattern  $X$ . If  $d_i(X) > d_j(X)$  for some  $i$  and for some  $j = 1, 2, \dots, m, i < j$ , the pattern  $X$  belongs to pattern class  $w_i$ . In other words, if the  $i$ th decision function,  $d_i(X)$ , has the largest value for a pattern  $X$ , then  $X$  belongs in  $w_i$ . Such an automatic classification scheme using a decision-making process is illustrated conceptually in Figure 2.5.

The decision functions can be generated in many ways, depending on a priori knowledge about the patterns to be recognized. Bayes classification rule can be used if the probability density distribution functions are known [Fu 76b].

Some of the more commonly used decision functions are:

(1). Linear discriminant function,

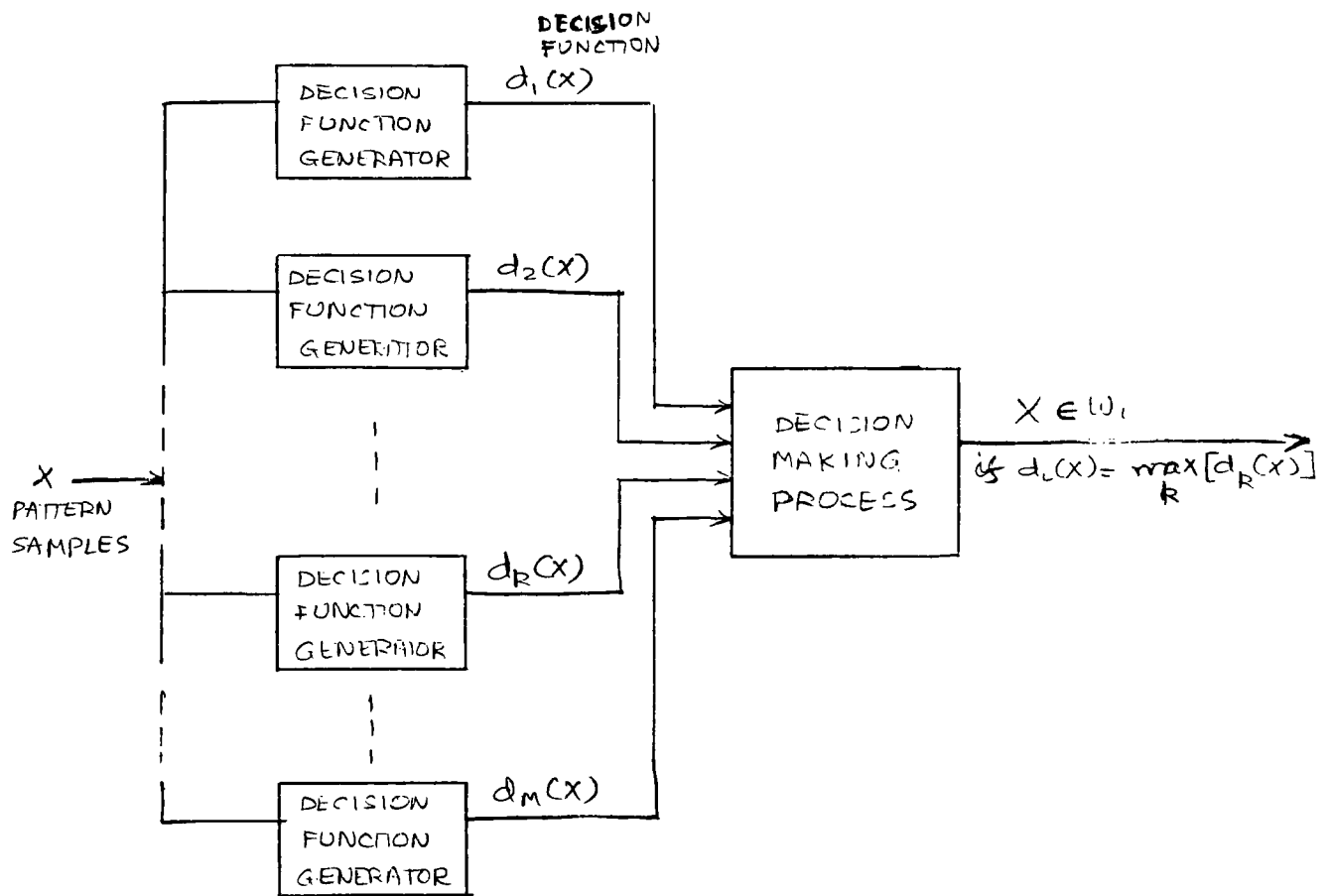


Fig 2-5 BLOCK DIAGRAM OF A PATTERN CLASSIFIER.

- (2). minimum distance classifier,
- (3). piecewise linear discriminant functions
- (Nearest neighbor)
- (4). polynomial discriminant functions.

Another method of pattern classification is to use the concept of distance functions [Tou 74]. One way of establishing a measure of similarity between pattern vectors, is by determining their proximity. For example, as shown in Figure 2.6, it is intuitively obvious that X belongs to pattern class  $w_i$ , solely on the basis that it is closer to the patterns of this class. However, it might be difficult to classify Y into either pattern class based on a measure of the proximity of this pattern to a class. The method of pattern classification by distance functions works better when the pattern classes tend to have clustering properties. Some of the distance functions used are:

- minimum distance method,
- maximin distance method,
- K-means algorithm,
- Isodata algorithm etc.

A similarity measure or dissimilarity measure gives a numerical value to the notion of closeness or distance between two objects. The choice of a similarity function depends on the ease of computing and its capacity for discrimination of various objects in the measurement space.

some of the quantitative measures are due to [Gid 76]:

Minkowsky metric:

$$d_1(F, T) = \left[ \sum_{j=1}^n |F_j - T_j|^{\frac{1}{\lambda}} \right]^{\lambda}$$

Camberra metric:

$$d_2(F, T) = \sum_{j=1}^n \frac{|F_j - T_j|}{|F_j + T_j|}$$

Chebychev metric:

$$d_3(F, T) = \max_j |F_j - T_j|$$

Quadratic metric:

$$d_4(F, T) = (F - T)^t Q (F - T)$$

where  $Q$  is a  $n \times n$  positive definite matrix.

Mahalanobis metric:

$$d_5(F, T) = (\det W)^{\frac{1}{p}} (F - T)^t W^{-1} (F - T)$$

Correlation

$$d_6(F, T) = \frac{\sum_{j=1}^n (F_j - \bar{F}_j)(T_j - \bar{T}_j)}{\left[ \sum_{j=1}^n (F_j - \bar{F}_j)^2 \sum_{j=1}^n (T_j - \bar{T}_j)^2 \right]^{1/2}}$$

"City block" metric:

$$d_7(F, T) = \sum_{j=1}^n w_j |F_j - T_j|$$

where  $w$  is a weight vector.

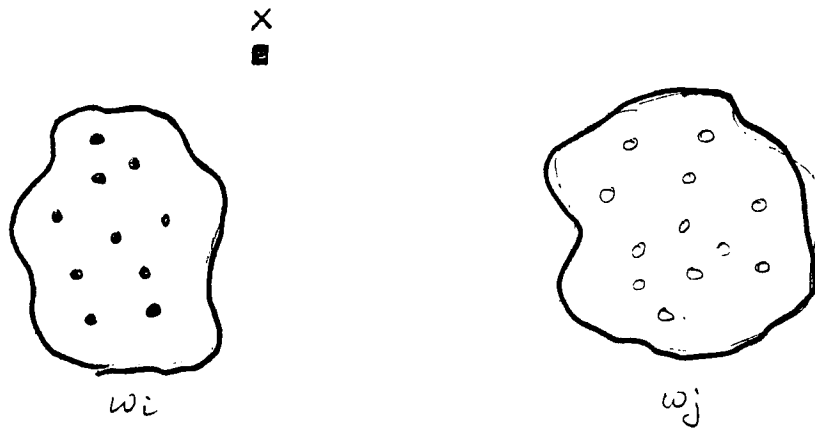
"Chi-square" metric:

$$d_8(F, T) = \sum_{j=1}^n \frac{1}{F_{\cdot j}} \left[ \frac{F_j}{F_{\cdot}} - \frac{T_j}{T_{\cdot}} \right]^2$$

where

$$F_{\cdot j} = \sum_{i=1}^m F_{ij} \text{ and } F_{\cdot} = \sum_{j=1}^n F_{\cdot j}.$$

$F$  and  $T$  are the objects.



PATTERNS CLASSIFIABLE BY PROXIMITY CONCEPT.

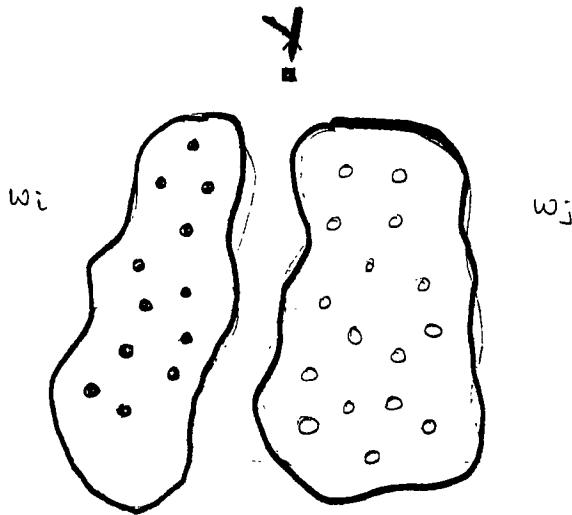


Fig 2-6 : PATTERNS NOT CLASSIFIABLE BY PROXIMITY CONCEPT

### 2.2.3 Learning -

With the linear classification functions described in section 2.2.2, perfect recognition is possible with correct values for the coefficients or weights. However, in practice, these values for the weights are usually not available. Under such circumstances, the pattern classifier can have the capability of estimating the best values for the weights from the input patterns. The basic idea is that, by observing patterns with known classifications, the classifier can automatically adjust the weights to achieve correct recognition. The performance of the classifier is supposed to improve as more and more sample patterns are applied. This process is called learning or training.

For each characteristic of the pattern, an initial weight is assigned. This weight is increased if its characteristic lead to correct recognition and decreased otherwise. The weight adjustment can be done by a fixed increment or by an appropriate fractional correlation rule [Fu 76b].

In the early pattern recognition system used by Uhr [Uhr 63], the recognition procedure utilized weights or amplifiers at various levels. The recognition procedure involved taking the difference between each characteristic of the input pattern and the standard dictionary pattern. These differences were then weighted by the corresponding pattern amplifiers, and then weighted again by general

amplifiers representing the average of pattern amplifiers across all patterns, which finally produced a weighted average difference between the input and the dictionary patterns. This average difference was multiplied by a final average difference amplifier to obtain a difference score. The pattern with the lowest difference score is selected as the correct standard pattern.

After a pattern was recognized, the pattern amplifiers in those patterns that had difference scores less than or only slightly above the difference score for the correct pattern, are modified. The correct pattern was compared with each of the similar patterns in turn. Each characteristic was examined individually, and a determination made as to whether the correct pattern would have been chosen if the choice had been made on the basis of this characteristic alone. If this one characteristic would have identified the correct pattern, then the corresponding amplifier was turned up by one. If it would have identified the wrong pattern then the amplifier was turned down by one. If a pattern had a higher difference score than the correct pattern, then the amplifiers were adjusted only in the incorrect pattern. Otherwise, amplifiers were adjusted in both patterns.

A simplified method of adjusting the weight factors was used in this thesis and is described in Chapter 3 in detail.



### 2.3 Syntactic Methods In Pattern Recognition

In some pattern recognition problems, the structural information that describes each pattern is important. In the case of picture recognition or scene analysis, the patterns being classified are usually quite complex and require a large number of features, hence the number of descriptors will be large. It becomes impossible to regard each descriptor as defining a class. Consequently, the requirements of recognition can be satisfied only by a description for each pattern rather than by the simple task of classification. Some language scripts like Devanagari, Telugu and Chinese, with thousands of characters, fall into this category.

One method of representing the hierarchical (tree-like) structural of each pattern is to describe the pattern in terms of simpler subpatterns, with each simpler subpattern again described in terms of even simpler subpatterns, etc. The simplest subpatterns selected, called "pattern primitives," should be much easier to recognize than the patterns themselves. The "language" providing the structural description of patterns in terms of a set of pattern primitives and their composition operations is sometimes called a "pattern description language." The rules governing the composition of primitives into patterns are usually specified by the "grammar" of the pattern description language. After each primitive within the pattern is identified, the recognition process is

accomplished by performing a syntax analysis or parsing of the "sentence" describing the given pattern to determine whether or not it is syntactically (or gramatically) correct with respect to the specified grammar. In addition, the syntax analysis also produces a structural description of the sentence representing the given pattern (usually in the form of a tree structure).

The various relations or composition operations defined among subpatterns usually can be expressed in terms of logical and/or mathematical operations. For example, if we choose "concatenation" as the only relation (composition operation) used in describing patterns, then for the pattern primitives shown in Figure 2.7, the rectangle shown in Figure 2.8 would be represented by the string 'aaabbbcccd'. More explicitly, if we use "+" for the head-to-tail concatenation operation, the rectangle in Figure 2.8 would be represented by the string 'a+a+a+b+b+b+c+c+c+d+d', and its corresponding tree-like structure would be as shown in Figure 2.9. Similarly, a slightly more complex example is given in Figure 2.10 using the pattern primitives given in Figure 2.7.

An alternative representation of a pattern's structural information is a "relational graph". In using a relational graph for pattern description, one can include any relation that can be conveniently determined from the pattern. Note that (1) concatenation is the only natural operation for one-dimensional languages, and (2) a graph can contain



Fig 2.7. Primitives.

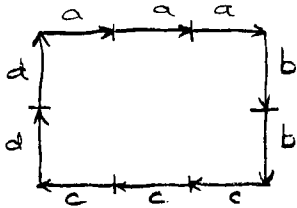


Fig 2.8. A Rectangle with primitives in Fig 2.7.

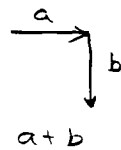
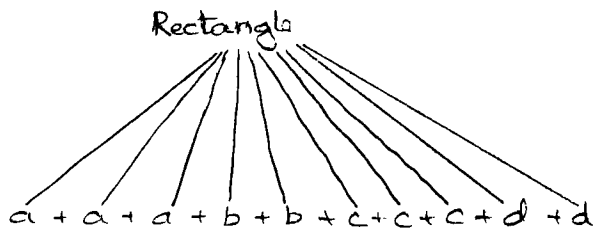


Fig 2.9. Tree structure for rectangle in Fig 2.8 using primitives in Fig 2.7

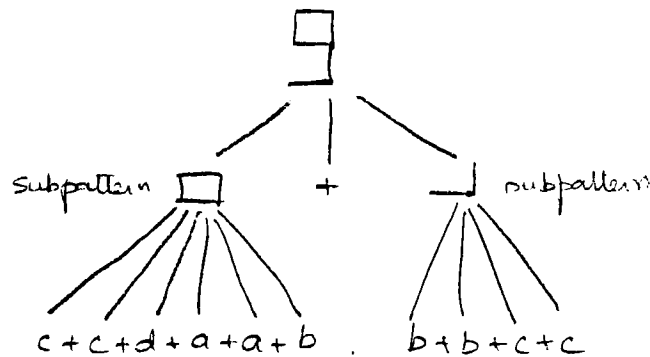
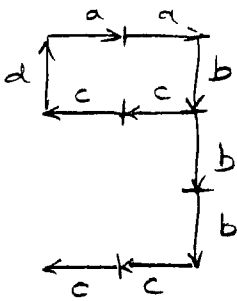
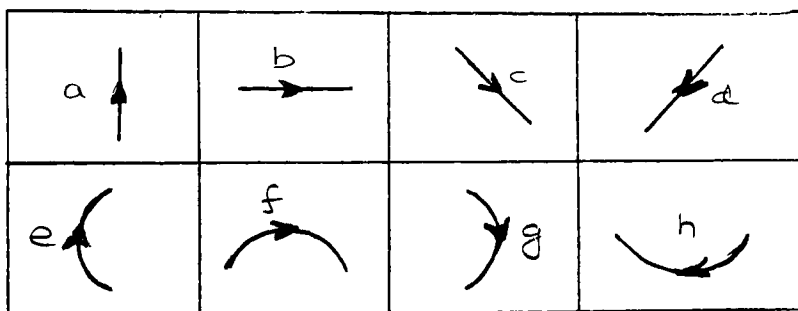


Fig 2.10 Pattern 9 and its structural description. [Fu 76]

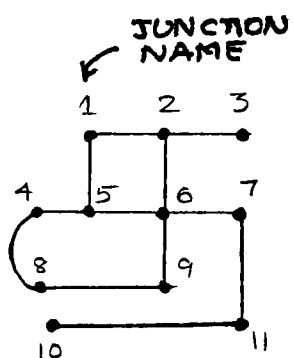
closed loops whereas a tree cannot, therefore, graphs can be used to express richer descriptions than trees. However, the use of tree structures does provide a direct way to adapt the techniques of formal language theory to the problem of compactly representing and analyzing patterns containing a significant structural content [Fu 76], whereas graphs do not provide a direct way of analyzing the structural content of a pattern.

A labeled graph approach was used to represent the structural and connectivity information of strokes in the Tamil script for handprinted character recognition by [Chi 80]. In this approach, the characters were assumed to be composed of line-like elements, called primitives (Figure 2.11a), satisfying certain relational constraints. Directed labeled graphs are used to describe the structural composition of characters in terms of their primitives and the relational constraints satisfied by them. For example, the nodes of the graph in Figure 2.11c correspond to the junction names of the letter 'த' in Figure 2.11b. A junction can either be an end point of a primitive or a point where two or more primitives meet. The label and direction of an edge represent the primitive joining the junctions named by the end nodes of the edge. For analysis, the graph of a letter is represented by a connectivity matrix. The recognition algorithm uses a topological matching procedure to compute and maximise the correlation coefficients. Basic symbols are first recognized to

# PRIMITIVES



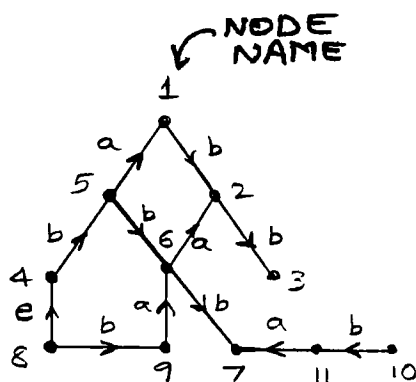
2.11 a




2.11 b CHARACTER  (TMA)

FIG 2.11

[chi 80]



2.11 c

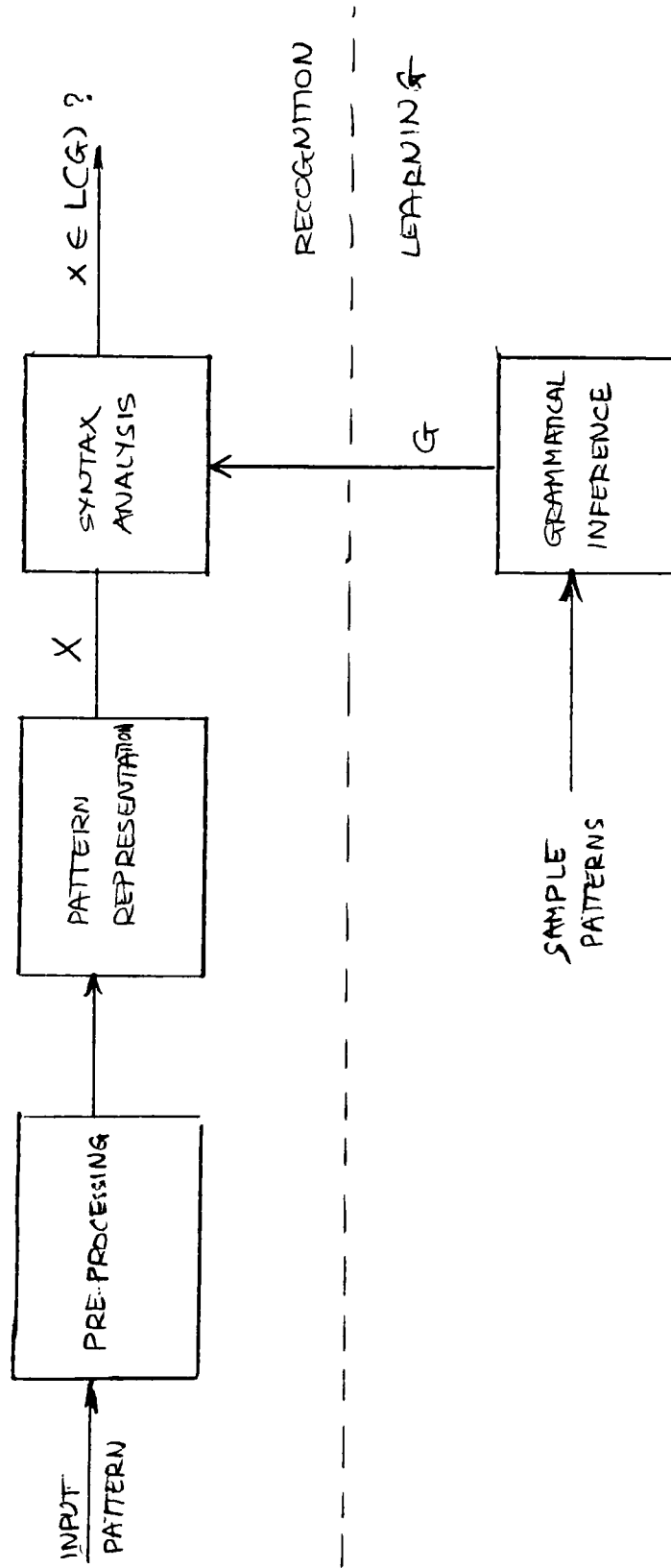
DIRECTED LABELED GRAPH OF 

identify derived symbols. An excellent recognition rate was obtained by using method.

### 2.3.1 Linguistic Pattern Recognition Systems -

A linguistic pattern recognition system general consists of three major phases: preprocessing, pattern description or representation, and syntax analysis, as shown in Figure 2.12. The functions of preprocessing include: (1) pattern encoding and approximation, and (2) filtering, restoration and enhancement.

In the preprocessing phase an input pattern is first coded into or approximated by some convenient form for further processing. For example, a black-and-white picture can be coded in terms of a grid or matrix of 0's and 1's, and a waveform can be approximated by its time samples or a truncated Fourier series expansion. Data compression is often applied at this stage to make the processing in the later stages of the system more efficient. Then, techniques of filtering, restoration and/or enhancement are used to clean up noise, to restore the degradation, and/or to improve the quality of the coded or approximated patterns. The output of the preprocessor, then would be patterns with reasonably "good quality". Each preprocessed pattern is then represented by a sentence-like structure, e.g., a string or a graph.



12  
 FIG 2.24 BLOCK DIAGRAM OF A LINGUISTIC PATTERN RECOGNITION SYSTEM.

The pattern representation process consists of: (1) pattern segmentation, and (2) primitive (feature) extraction. Each preprocessed pattern is segmented into subpatterns and pattern primitives based on prespecified syntactic or composition operations and, in turn, each subpattern is identified with a given set of pattern primitives. Each pattern is then represented by a set of primitives with specified syntactic operations. The decision on whether or not the representation (pattern) is syntactically correct (ie. belongs to the class of patterns described by the given syntax or grammar) will be performed by the syntax analyzer or "parser". The syntax analyzer usually can produce a complete syntactic description of the pattern in terms of a parse-tree, provided the pattern is syntactically correct. Otherwise, the pattern is either rejected or analyzed on the basis of other given grammars, which presumably describe other possible classes of patterns under consideration.

The simplest form of recognition is "template matching". The string of primitives representing the input pattern is matched against strings of primitives representing each reference pattern, and the input is classified in the same class as the reference pattern which is the "best" match. More complex recognition methods involve a complete parsing of the input string exploring the complete hierarchical structural description of the pattern. The selection of appropriate approach for recognition



usually depends on the problem requirement.

Obtaining a grammar that describes the structural information about the class of patterns under study requires a grammatical inference machine which infers a grammar from a given set of training patterns in language-like representations. This is analogous to the "learning" process in statistical pattern recognition systems. The structural description of the class of patterns under study is learned from the actual sample patterns from that class. The learned description, in the form of a grammar is then used for pattern description and syntax analysis.

Practical applications of linguistic pattern recognition include the recognition of English, Chinese [Cha 73], and Devanagari characters [Sin 79], spoken digits, mathematical expressions [And 68], the classification of bubble-chamber and spark-chamber photographs [Sna 68] [Bha 72], chromosomes and finger print images [Mca 75], and the identification of machine parts [Vam 73].

An excellent example of the use of the linguistic method is PLANG, a language describing Devanagari script [Sin 83]. This script consisted of characters with vertical, horizontal and cursive strokes. The patterns are described in a pattern description language for recognition and analysis.

A sentence in PLANG described a two-dimensional pattern in terms of primitives and composed macros specifying their relationships with respect to various regions of the picture frame or in terms of subpatterns operated by a frame function(s). A partitioning function was assumed which partitioned the picture into nine equal regions as shown in Figure 2.13. Every partitioned region could be partitioned further to get finer 'regions' by recursively applying the partition function on a partitioned region. The union of regions could be defined by using the U operator. A sentence in PLANG was represented as the outcome of a 'frame-function' operating upon a list of 'frames' or another PLANG sentence. Two or more picture frames could be combined in a desired manner to make a larger picture frame, or an original picture frame could be transformed to a new picture frame in a desired manner utilizing a set of frame-functions. Two frame functions named 'superimpose' (denoted by '\*') and 'append' (denoted by '.') were used for Devanagari script. HRZ (horizontal line), VERT (vertical line), LTD (left-going diagonal), RTD(right-going diagonal), CURVE (arbitrary curve passing through specified regions in that order), and DOT (a dot) are selected as the primitives, and were illustrated in Figure 2.13. Composition of the Devanagari character KRAIN and its PLANG description are given in Figure 2.14.

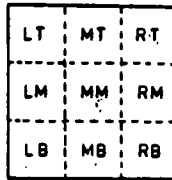


Fig. 1. Partitioning function.

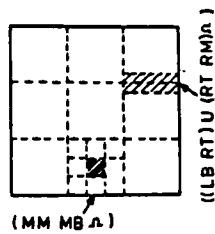


Fig. 2. Operator U and repartitioning.

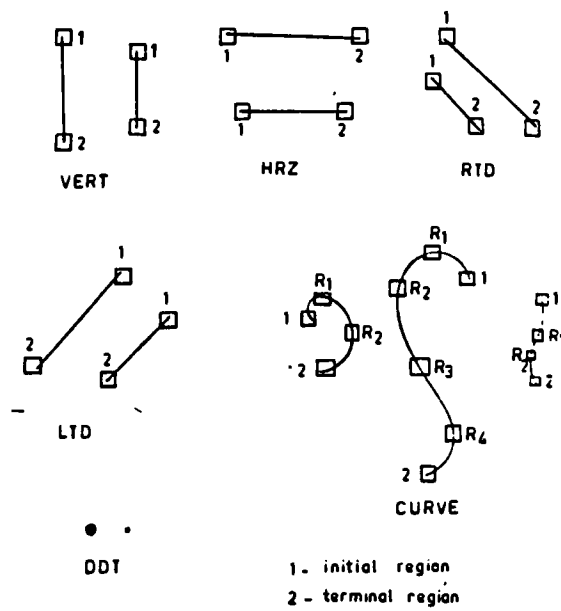
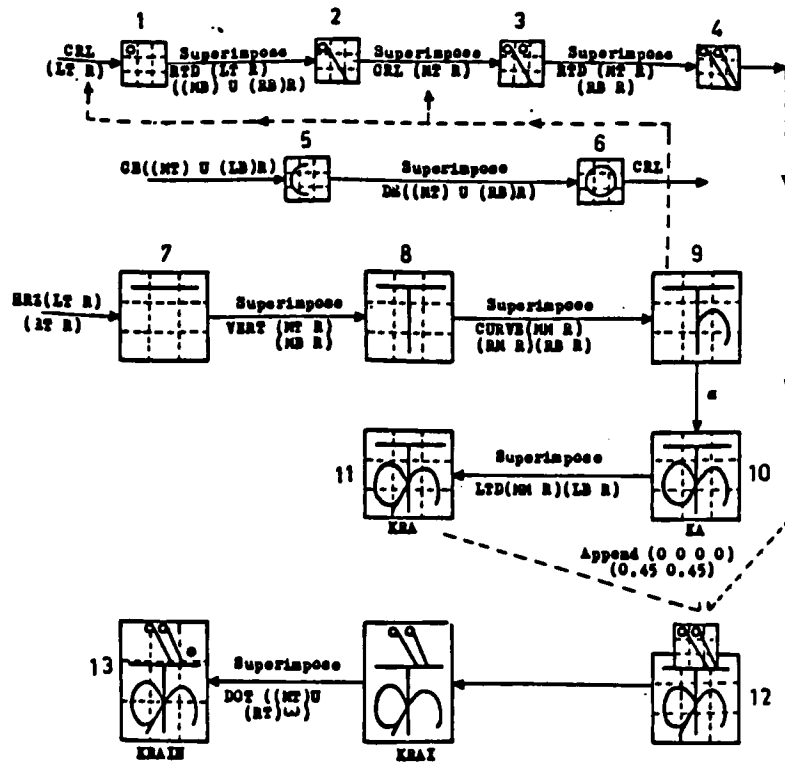


Fig. 3. Primitives.

2.13 BUILD PRIMITIVES IN PATTERN DESCRIPTION LANGUAGE PLANG.  
[Sin 83]



$\alpha$  : Superimpose CURVE(MM R)((MT LM) U (RT LM) R)((LM LM) U (MM LB) R)((MT MB) U (LB MM) R) (MM LB)

Composition of Devanagari composite character KRAIN.

```

DEVKRAIN Begin
Describe KRAIN ((KRAI)(DOT((MT) U (RT)Ω))))
Describe KRAI ((. (0 0 0 0) (0.45 0.45))((KRA ((LM) U (RB)Ω)) (AI ((LT) U (RT)Ω))))
Compose ((KRA(R))((KA(R))(LTD(MM R) (LB R))))
Compose ((KA (R)) ((HRZ(LT R)(RT R)) (VERT(MT R)(MB R))(CURVE (MM R)(RM R) (RB R)) (CURVE
((MT MM) U (LB MM) R) ((MT LM) U (RT LM) R) ((LM LM) U (MM LB) R) (MM LB)
U (MT MB) R) ((MT MM) U (LB MM) R))))
Compose ((AI(R))((CRL (LT R))(RTD (LT R)((MB) U (RB) R))(CRL (MT R))(RTD (MT R) (RB R))))
Compose ((CRL (R))((CE((MT) U (LB) R))(DE(MT) U (RB) R))))
Compose ((CE(R))(CURVE (RT R) ((MT) U (RB LT) R) ((LM) U (MM) R)((MB) U (RT LB) R)(RB R))
Compose ((DE(R))(CURVE (LT R)((MT) U (LB RT) R) ((RM) U (MM) R)((MB) U (LT RB) R)(LB R))
Ω (coordinates of lower left corner of picture-frame height width)
End

```

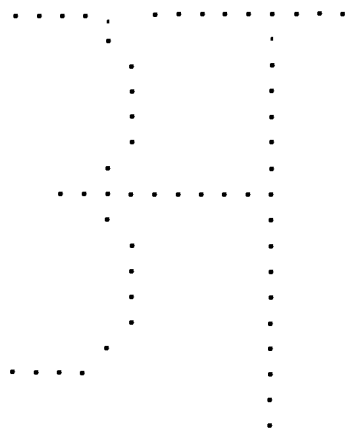
2.14 Composition and PLANG description of Devanagari character KRAIN. [Sin 83].

A PLANG description for all symbols of the script were stored. A local feature extraction operation was performed in which every point of the pattern was assigned a label depending upon the local property it exhibited with respect to its neighboring points. The local properties included whether the point in question formed the part of a vertical line, horizontal line, right-going diagonal, left-going diagonal or do not care. These labels were used by a goal oriented top-down parser for primitive recognition. In Figure 2.15, a hierarchic representation of the PLANG description for the symbol KA is shown. The parser looks for the frame function and looks for the appropriate primitives in the pattern. If it is successful, the parser moves on to other subtrees towards the right. If, at any stage, the goal is unsuccessful, a penalty is entered and if the penalty exceeds a preset threshold value, the parser picks up the description of the next probable symbol.

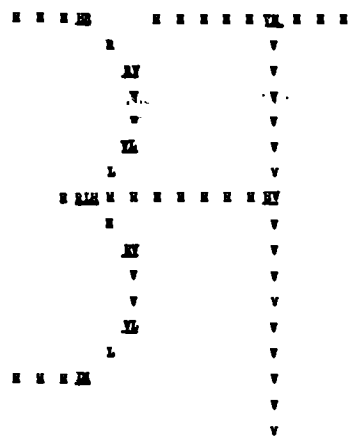
Contextual information about the script helped achieve a recognition rate of more than 90 percent.

### 2.3.2 Selection Of Pattern Primitives -

The determination of primitives with which the pattern of interest may be described is largely influenced by the nature of the data; there is no general solution for the primitive selection problem [Fu 76]. The primitives should serve as basic pattern elements to provide a compact but adequate description of the data in terms of the specified



(a)



(b)

Fig. 6. (a) Thinned pattern. (b) Labeled pattern.

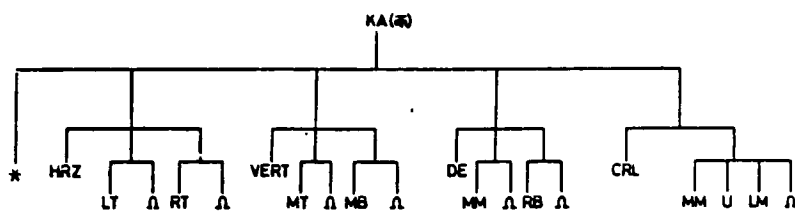


Fig. 8. Parsing tree for symbol "KA".

2.15 HIERARCHIC REPRESENTATION FOR SYMBOL 'KA' [Sin 83]

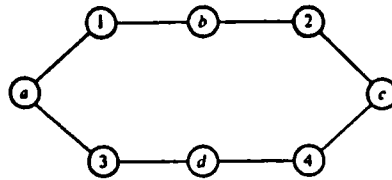
structural relations (e.g., the concatenation relation). The primitives should be easily extracted or recognized by existing non-linguistic methods, since they are considered to be simple and compact patterns and their structural information is not important. For example, for speech patterns, phonemes are naturally considered as a "good" set of primitives with the concatenation relation [Cho 63]. Similarly, strokes have been suggested as primitives in describing handwriting [Ede 61]. However, for general pictorial patterns, there is no such universal picture element.

Freeman's chain code [Fre 61][Fre 62] is a common set of primitives used to describe the boundaries or skeletons in a pattern. In this scheme, a rectangular grid is overlaid on a two-dimensional pattern, and the straight line segments are used to connect the grid points falling closest to the pattern. Each line segment is assigned an octal digit according to its slope. Patterns are thus represented by chains or strings of octal digits. Figure 2.16 illustrates the primitives and the coded string describing a curve. Simple manipulations like rotation, expansion, measurement of curve length, and determination of pattern self-intersections can be easily carried out. Knoke and Wiley [Kno 67] and Feder [Fed 68] used this method for describing hand-printed characters.

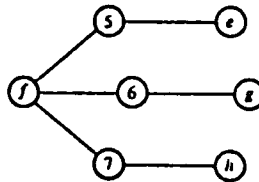
A set of primitives encoding geometric patterns in terms of regions has been proposed by [Pav 68]. In this case, the basic primitives are halfplanes in the pattern space. It can be shown that any figure (or arbitrary polygon) may be expressed as the union of a finite number of convex polygons. Each convex polygon can, in turn, be represented as the intersection of a finite number of halfplanes. By defining a suitable ordering (a sequence) of the convex polygons composing the arbitrary polygon, it is possible to determine a unique minimal set of maximal polygons, called a primary subset, the union of which is the given polygon. As a linguistic analogy, a figure can be thought of as a 'sentence', the convex polygons composing it as 'words,' and the halfplanes as 'letters'. A more general selection procedure of pattern primitives based on regions was proposed by Rosenfield and Strong [Ros 71].

Another form of representing polygonal figures is the use of primary graphs [Pav1 72][Pav2 72]. The primary graph of a polygon A is one whose nodes correspond to the nuclei and primary subsets of A, and whose branches connect each nucleus to all the primary subsets containing it. An example is given in Figure 2.17. Primary subsets and nuclei of polygons approximating the figures are shown in Figure 2.17a (shaded areas are nuclei). Primary graphs for the corresponding polygons in Figure 2.17a are given in Figure 2.17b. This form of representation provides information about the topology of the picture. Also, patterns

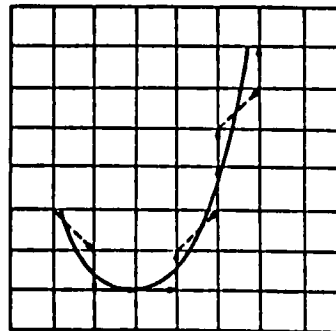




A diagram showing a stepped profile with five segments labeled  $e$ ,  $f$ ,  $g$ ,  $f$ , and  $h$  from top to bottom. The left side of the profile is a vertical line with three shaded rectangular blocks at the top, middle, and bottom. The segments  $e$ ,  $g$ , and  $h$  are horizontal, while  $f$  and  $f$  are vertical.



A square is divided into eight triangles by its diagonals and lines connecting the midpoints of opposite sides. The triangles are numbered 1 through 7, with the central triangle being the eighth triangle.



## 2.16 Freeman's chain code [Fre 62]

represented by graphs can be formally described by graph grammars which can be analysed.

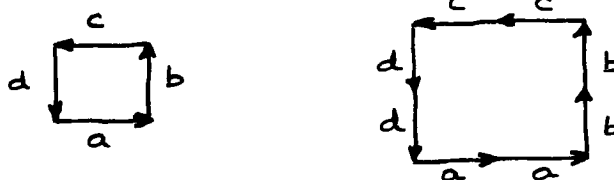
### 2.3.3 Pattern Grammar -

Once the pattern primitives are selected, the next step is to construct a grammar that will generate a language to describe the patterns under study. We will consider two classes of languages, namely, finite-state versus context-free and context-sensitive. Finite state automata, which are easy to implement recognize or accept finite-state languages. However the descriptive power of finite-state languages is weaker than that of context-free and context-sensitive languages which require non-deterministic parsing procedures. A precedence language may be used for pattern description in order to obtain efficient analysis. On the other hand, a context-free programmed grammar generating a context-sensitive language may be selected in order to describe the patterns effectively.

As an example consider the language

$$L = \{a^n b^n c^n d^n \mid n \geq 1\}$$

which could be interpreted as the language describing squares of side length  $n = 1, 2, \dots$

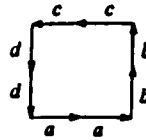


$L$  is a context sensitive language and can be generated by a context-sensitive grammar or a context-free programmed grammar as shown in Figure 2.18.

# The language

$$L = \{a^n b^n c^n d^n | n \geq 1\}$$

could be interpreted as the language describing squares of side length  $n = 1, 2, \dots$



$L$  is known as a context-sensitive language, and can be generated in the following two ways.

1) A context-sensitive grammar

$$G_1 = (V_N, V_T, P, S)$$

where

$$V_N = \{S, A, B, C, D, E, F, G\}$$

$$V_T = \{a, b, c, d\}$$

$$P: \begin{array}{ll} S \rightarrow aAb & dG \rightarrow Gd \\ A \rightarrow aAC & aG \rightarrow abcD \\ A \rightarrow D & bG \rightarrow bbcD \\ Dc \rightarrow cD & dFB \rightarrow dFd \\ Dd \rightarrow dD & dFd \rightarrow Fdd \\ DC \rightarrow EC & cF \rightarrow Fc \\ EC \rightarrow Ed & bF \rightarrow bbc \\ DB \rightarrow FB & aF \rightarrow ab \\ Ed \rightarrow Gd & bB \rightarrow bcd \\ cG \rightarrow Gc. \end{array}$$

FIG. 2.18a Context-sensitive grammar [Fu 76].

A context-free programmed grammar

$$G_2 = (V_N, V_T, P, S, J)$$

where

$$V_N = \{S, A, B, C, D\}$$

$$V_T = \{a, b, c, d\}$$

$$J = \{1, 2, 3, 4, 5, 6, 7\}$$

P:	Label	Core	Success field	Failure field
1	$S \rightarrow aAB$		$\{2, 3\}$	$\{\emptyset\}$
2	$A \rightarrow aAC$		$\{2, 3\}$	$\{\emptyset\}$
3	$A \rightarrow D$		$\{4\}$	$\{\emptyset\}$
4	$C \rightarrow d$		$\{5\}$	$\{6\}$
5	$D \rightarrow bDc$		$\{4\}$	$\{\emptyset\}$
6	$B \rightarrow d$		$\{7\}$	$\{\emptyset\}$
7	$D \rightarrow bc$		$\{\emptyset\}$	$\{\emptyset\}$

FIG. 2.18 b. Context-free Programmed grammar. [Fu 76]

Although many classes of patterns intuitively appear to be context-sensitive, context-sensitive grammars have rarely been used for pattern description simply because of their complexity. Context-free languages have been used to describe patterns such as English characters, chromosome images, spark-chamber pictures [Sha 68], chemical structures, finger-print patterns and spoken digits [And 68]. Figure 2.19a describes a context-free grammar describing the chromosome images shown in Figure 2.19b.

In patterns using string grammars, the only relation between subpatterns and/or primitives is concatenation; that is, each subpattern or primitive can be connected only at the left or right. This one-dimensional relation is not very effective in describing two- or three-dimensional patterns like mathematical expressions, pictograms etc. An attempt was made to include more useful relations by [Nar 70]. As an example, TRIANGLE(a,b,c) means that a ternary relation TRIANGLE, is satisfied by the line segments a, b and c, and ABOVE(X,Y) means that X is above Y. Similarly the mathematical expression

$$\frac{a+b}{c}$$

can be described by

ABOVE(ABOVE(LEFT(a,LEFT(+,b)),--),c)

where LEFT(X,Y) means that X is to the left of Y. A grammar describing the houses in Figure 2.20a is shown in Figure 2.20b.

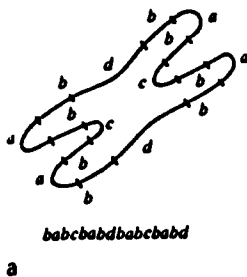
$V_N = \{ \langle \text{submedian chromosome} \rangle, \langle \text{telocentric chromosome} \rangle, \langle \text{arm pair} \rangle, \langle \text{left part} \rangle, \langle \text{right part} \rangle, \langle \text{arm} \rangle, \langle \text{side} \rangle, \langle \text{bottom} \rangle \}$

$$V_T = \left\{ \begin{array}{c} \hat{\cap} \\ a \end{array} , \begin{array}{c} | \\ b \end{array} , \begin{array}{c} \cup \\ c \end{array} , \begin{array}{c} \} \\ d \end{array} , \begin{array}{c} \smile \\ e \end{array} \right\}$$

and

$P$ :  $\langle \text{submedian chromosome} \rangle \rightarrow \langle \text{arm pair} \rangle \langle \text{arm pair} \rangle$   
 $\langle \text{telocentric chromosome} \rangle \rightarrow \langle \text{bottom} \rangle \langle \text{arm pair} \rangle$   
 $\langle \text{arm pair} \rangle \rightarrow \langle \text{side} \rangle \langle \text{arm pair} \rangle$   
 $\langle \text{arm pair} \rangle \rightarrow \langle \text{arm pair} \rangle \langle \text{side} \rangle$   
 $\langle \text{arm pair} \rangle \rightarrow \langle \text{arm} \rangle \langle \text{right part} \rangle$   
 $\langle \text{arm pair} \rangle \rightarrow \langle \text{left part} \rangle \langle \text{arm} \rangle$   
 $\langle \text{left part} \rangle \rightarrow \langle \text{arm} \rangle c$   
 $\langle \text{right part} \rangle \rightarrow c \langle \text{arm} \rangle$   
 $\langle \text{bottom} \rangle \rightarrow b \langle \text{bottom} \rangle$   
 $\langle \text{bottom} \rangle \rightarrow \langle \text{bottom} \rangle b$   
 $\langle \text{bottom} \rangle \rightarrow e$   
 $\langle \text{side} \rangle \rightarrow b \langle \text{side} \rangle$   
 $\langle \text{side} \rangle \rightarrow \langle \text{side} \rangle b$   
 $\langle \text{side} \rangle \rightarrow b$   
 $\langle \text{side} \rangle \rightarrow d$   
 $\langle \text{arm} \rangle \rightarrow b \langle \text{arm} \rangle$   
 $\langle \text{arm} \rangle \rightarrow \langle \text{arm} \rangle b$   
 $\langle \text{arm} \rangle \rightarrow a$

2.19a context-free grammar describing chromosome images described in 2.19b. [Fu 76].

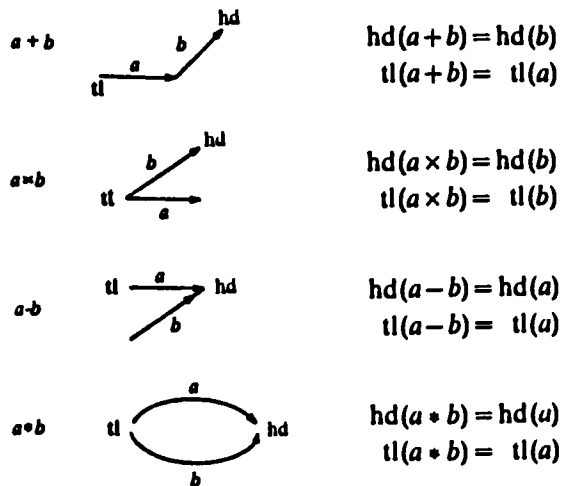


2.19b a) Submedian chromosome and b) telocentric chromosome

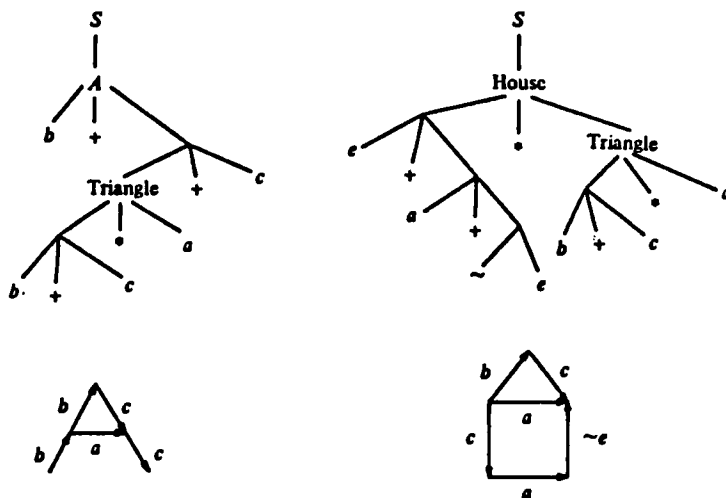
[Sha 70], by attaching a 'head' (hd) and a 'tail' (tl) to each primitive, has used the four binary operators  $+$ ,  $\times$ ,  $-$ , and  $*$  for defining binary concatenation relations between primitives (Figure 2.21). Each pictorial pattern can be represented by a 'labeled branch-oriented graph' or relational graph, where nodes represent subpatterns or primitives and branches denote (binary) relations (Figure 2.22). Feder [Fed 71] has formalized a 'plex' grammar which generates languages with terminals having an arbitrary number of attaching points for connecting to other primitives or subpatterns. Pfaltz and Rosenfield have extended the concept of string grammars to grammars for labeled graphs called 'webs' [Pfa 71]. Each production describes the rewriting of a graph A into another graph B and also contains an 'embedding' rule E which specifies the connection of B to its surrounding graph (host web) when A is rewritten. Here, the terminals or primitives are represented as vertices in the graph.

#### 2.3.4 Syntax Analysis As Recognition Procedure -

After selecting the appropriate pattern primitives and their suitable pattern grammar representing the concatenation relations of the primitives, the next step in syntactic pattern recognition method would be to perform syntax analysis. Syntax analysis or parsing is necessary if a complete description of the input pattern is required for recognition. As mentioned above, finite-state automata



2.21 Concatenation relations between primitives [Fu76]



PDL structural description of "A" and "House"

Fig 2.22. Patterns represented by labeled branch-oriented graphs. [Fu76].



in fig 2.20b.

$$G = (V_N, V_T, P, S)$$

where

$$V_N = \{ \langle \text{house} \rangle, \langle \text{side view} \rangle, \langle \text{front view} \rangle, \langle \text{roof} \rangle, \langle \text{gable} \rangle, \langle \text{wall} \rangle, \langle \text{chimney} \rangle, \langle \text{windows} \rangle, \langle \text{door} \rangle \}$$

$$V_T = \left\{ \square, \sqsupset, \boxplus, \triangle, \square, \nabla, \rightarrow, (, ), \odot, \ominus, \uparrow, \mapsto \right\}$$

$$S = \langle \text{house} \rangle$$

$$P: \langle \text{door} \rangle \rightarrow \boxplus$$

$$\langle \text{window} \rangle \rightarrow \boxplus, \langle \text{windows} \rangle \rightarrow \rightarrow (\langle \text{windows} \rangle, \boxplus)$$

$$\langle \text{chimney} \rangle \rightarrow \sqsupset, \langle \text{chimney} \rangle \rightarrow \sqsupset$$

$$\langle \text{wall} \rangle \rightarrow \square, \langle \text{wall} \rangle \rightarrow \ominus (\langle \text{door} \rangle, \square)$$

$$\langle \text{wall} \rangle \rightarrow \odot (\langle \text{windows} \rangle, \square)$$

$$\langle \text{gable} \rangle \rightarrow \triangle, \langle \text{gable} \rangle \rightarrow \uparrow (\langle \text{chimney} \rangle, \triangle)$$

$$\langle \text{roof} \rangle \rightarrow \nabla, \langle \text{roof} \rangle \rightarrow \uparrow (\langle \text{chimney} \rangle, \nabla)$$

$$\langle \text{front view} \rangle \rightarrow \uparrow (\langle \text{gable} \rangle, \langle \text{wall} \rangle)$$

$$\langle \text{side view} \rangle \rightarrow \uparrow (\langle \text{roof} \rangle, \langle \text{wall} \rangle)$$

$$\langle \text{house} \rangle \rightarrow \langle \text{front view} \rangle$$

$$\langle \text{house} \rangle \rightarrow \mapsto (\langle \text{house} \rangle, \langle \text{side view} \rangle)$$

The notation

$\rightarrow (X, Y)$  means that  $X$  is to the right of  $Y$ ,

$\odot (X, Y)$  means that  $X$  is inside of  $Y$ ,

$\ominus (X, Y)$  means that  $X$  is inside on the bottom of  $Y$ ,

$\uparrow (X, Y)$  means that  $X$  rests on top of  $Y$ ,

$\mapsto (X, Y)$  means that  $X$  rests to the right of  $Y$ .

Fig. 2.20a. Grammar describing houses in 2.20b. [Fu 76].




House	Description
	$\uparrow (\triangle, \square)$
	$\uparrow (\uparrow (\sqsupset, \triangle), \ominus (\boxplus, \square))$
	$\mapsto (\uparrow (\uparrow (\square, \nabla), \odot (\boxplus, \square)), \uparrow (\triangle, \ominus (\boxplus, \square)))$

Fig. 2.20b. Pattern description of houses. [Fu 76]

recognize or accept finite-state languages. If a class of patterns can be described by a finite-state language, a finite-state automaton can then be constructed to recognize the strings or sentences describing this class of patterns.

When a context-free language is used to describe a class of patterns, the corresponding recognition device is, in general, a non-deterministic pushdown automaton. The output of the syntax analyzer usually includes not only the decision of accepting the string generated by the given grammar, but also the derivation tree of the string, which in turn, gives the complete structural description of the pattern.

## 2.4 Descriptive Characteristics Of Telugu Characters

The telugu alphabet, with more than 2000 characters, consists of vowels, consonants, and consonant vowel combinations (C-V letters), and combinations of consonants and/or C-V letters (conjunct consonants). There are 16 vowels (Figure 2.23a) and 36 consonants (Figure 2.23b). A vowel following a consonant takes on a different graphic form called a 'vowel sign' [raj 77]. Vowel signs corresponding to the vowels in Figure 2.23a are shown in Figure 2.23c. The vowels and consonants combine to form 576 different C-V letters. The C-V letters are formed by adding vowel signs to consonants at appropriate places. Figure 2.23d shows some C-V letters obtained by combining the first consonant in Figure 2.23b with all the vowels. Similarly

అ ఆ ఇ ఈ ఉ ఊ ఋ  
ఎ ఏ ఐ ఒ ఓ ఔ అం అః

FIG. 1. Vowels.

Fig. 2.23a Vowels

క ఖ గ ఘ జ చ ఛ జ  
ఝ ఞ ట ఠ డ ఢ ణ త  
థ ద ధ న ప ఫ బ భ  
మ య ర ల శ వ శ.  
ష స హ క్ష

FIG. 2. Consonants.

Fig. 2.23b. Consonants.

✓ → అ ఐ ఈ ఇ ఉ ఊ ఋ  
→ క ఖ గ ఘ జ చ ఛ జ  
ఝ ఞ ట ఠ డ ఢ ణ త  
థ ద ధ న ప ఫ బ భ  
మ య ర ల శ వ శ.  
ష స హ క్ష

FIG. 3. Vowel signs (starred vowel signs occur only with the consonants 3, 9, 12, 17, 18, 19, 26, and 27).

Fig. 2.23c. Vowel signs corresponding to vowels in fig 2.23a

క కా కి కీ కు కూ కృ కృ  
కై కే కై కొ కో కౌ కం కః

2.23 d.

ఝ ఝ ఝ ఝ ఝ ఝ  
ఝ ఝ ఝ ఝ ఝ ఝ  
ఝ ఝ ఝ ఝ ఝ ఝ

2.23 e.

Fig 2.23d & e. Examples of C-V letters. [Raj 77]

the characters shown in Figure 2.23e are obtained using the ninth consonant. It is clear from the above description that vowels, consonants, and C-V letters can be realized by superimposing shapes called "build-primitives" at appropriate places over certain "basic-letters". The basic letters and build-primitives are shown in Figure 2.23f and 2.23g, respectively.

C-V letters and consonants may be combined in many different ways to produce thousands of different characters called 'conjunct-constants'. In general, one C-V letter and one or more of the consonants combine to form a conjunct consonant. Conjunct consonants with more than one consonant are very rare. If such cases are excluded, a conjunct consonant will consist of two components. The first component, called the "main component", is a C-V letter. The shape and location of the second component depends on which consonant it is. For some consonants, special symbols, called 'conjunct-primitives' (Figure 2.23h), are substituted to form the second component (Figure 2.23i). In some cases, the corresponding consonant (without the first vowel sign) is written below the main component; some examples are illustrated in Figure 2.23j.

## 2.5 Previous Work In Recognition Of Some Indian Scripts.

Kajasekharan and Deekshatulu [Kaj 77] have successfully used structural methods for the recognition of Telugu characters. Considerable work was done in recognition of

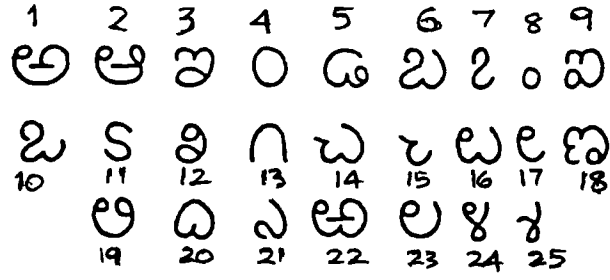


Fig. 2.23f. Basic Letters.

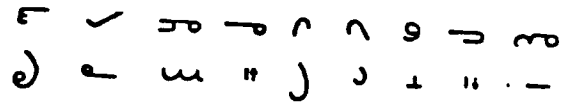


Fig. 2.23g. Build-primitives.

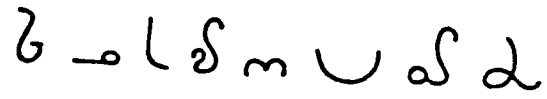


Fig. 2.23h Conjoin-primitives.

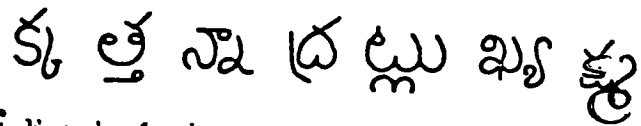


Fig. 2.23i. Examples of conjoin consonants with conjoin-primitives.



Fig. 2.23j. Examples of conjoin consonants in which consonants without the first vowel sign appear below C V letters. [Raj 77].

other Indian language scripts like Devanagari, Bengali and Tamil. Siromoney and Chandrasekharan [Sir 78] used a condensed run method for recognizing printed Tamil characters. A labelled graph representation for Tamil characters was used by Chinnuswamy [Chi 80]. Sinha [Sin 79] used syntactic approach and a picture language to describe the build primitives of Devanagari characters. A knowledge based system was developed later on to recognize words in the Devanagari script [Sin 84]. A generalized formal approach for description and analysis of major Indian languages was proposed by Datta [Dat 85]. A combination of structural and statistical methods for recognition of Tamil, Malayalam and Devanagari scripts was attempted by Chandrasekharan [Cha 85]. Ray and Chatterjee [Ray 85] designed a classifier for Bengali characters based on nearest neighborhood technique. A character recognition system, based on the contextual information in composing the Tamil characters, was developed by Chandrasekharan [Cha 83]. A heirarchical decision tree classification scheme is used by Marwah et al [Mar 84], based on the fact that characters in Devanagari script can be constructed using certain basic primitives. (Note: Bengali, Malayalam, Tamil and Telugu are regional languages in India and have their own distinctive scripts. Devanagari script is used for languages like Hindi, Marathi, Sanskrit etc.)

#### 2.5.1 Recognition Of Telugu Characters Using Structural Methods -

After a careful analysis of the Telugu characters, it can be found that all the possible Telugu characters can be realized by superimposing certain primitive shapes over 25 basic letters. Based on this observation Rajasekharan et al [Raj 77] devised a two-stage recognition system. In the first stage, a given pattern was examined for the presence of primitives. If they were found, they were removed from the given pattern after their presence was noted. The primitives were looked for in the character in a definite order since some primitives occlude others. Also, the search for a particular primitive was limited to a particular region of the pattern. After the removal of the primitive shapes, the remaining basic letter was recognized using an on-the-line coding procedure. Overall recognition of the given characters was accomplished by using a decision tree which made use of the primitives and basic letters that were present in the input pattern, and contextual information for composing the characters.

A block diagram in Figure 2.24 describes various subsystems of the Telugu character recognition system. The digitized images were segmented to separate composite characters into basic character primitives. A histogram of the number of figure points present in each row of the input pattern was used for this segmentation. In a binary pattern of 0's and 1's, figure points were represented by 1's and background points by 0's.

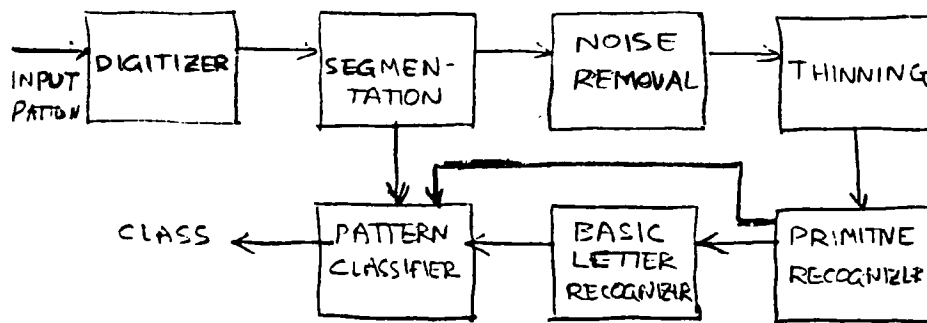


FIG 2.24 TELOGU CHARACTER RECOGNITION SYSTEM [Raj 77]

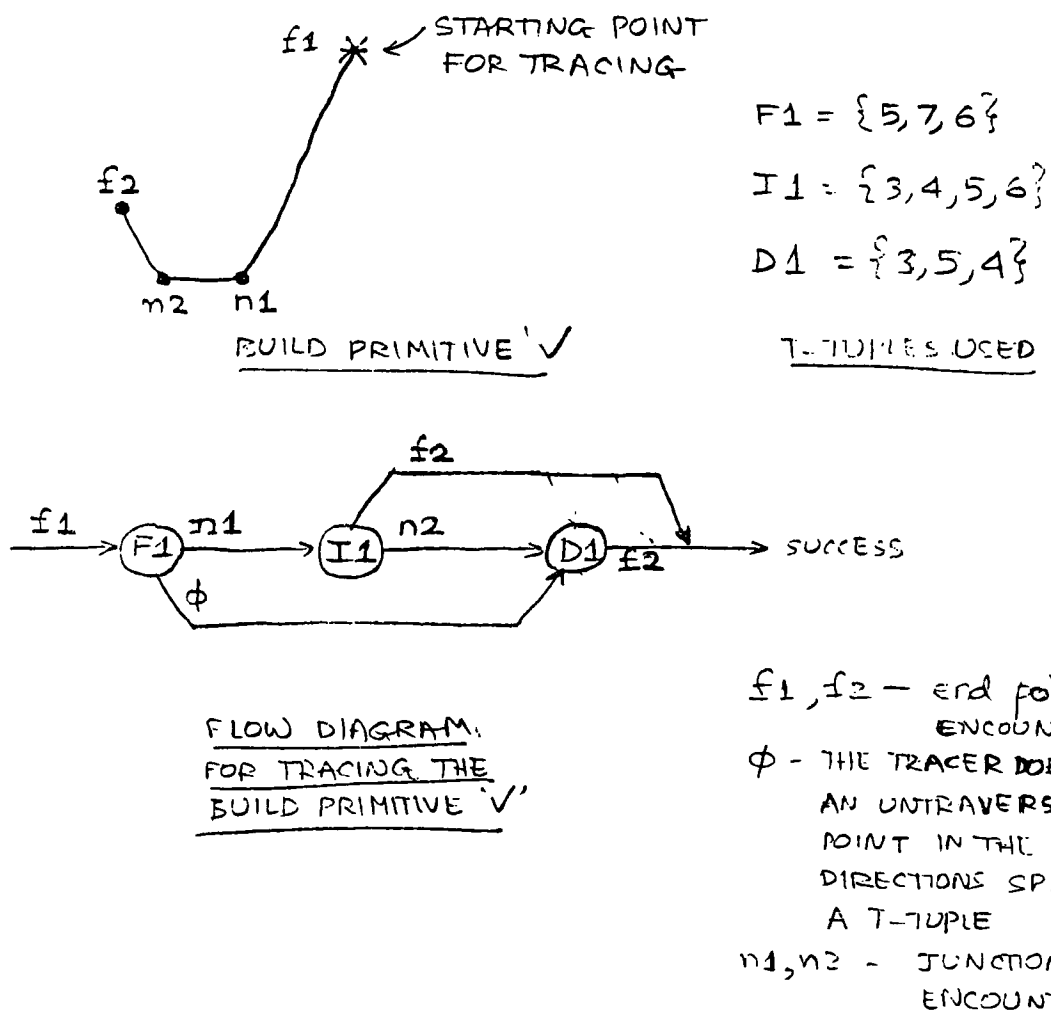
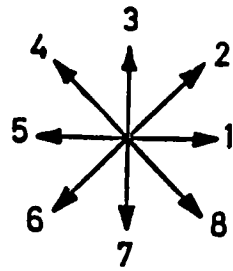


FIG 2.27 TELOGU BUILD PRIMITIVE 'V' AND ITS T-TUPLES. [Raj 77]



Noise was removed in two phases. In the first phase, if the number of figure points in the 8-neighbourhood (Figure 2.25) of a background point was greater than 5, the corresponding point was filled with a figure point. In the second phase, if the number of figure points in the 8-neighbourhood of a figure point was less than 4, then the corresponding point was removed. The thinning algorithm developed by Deutsch [Deu 72], was then used to extract the skeleton of the image.

The sequential template matching procedure makes use of a directed curve-tracer which traces a thinned-line pattern by choosing successive points in one of the allowed directions. The set of allowed directions is an ordered subset of the set of eight principal directions (Figure 2.25). Let the eight principal directions be denoted by  $N_x = \{1, 2, \dots, 8\}$ . Let  $R = \{R_1, R_2, \dots, R_T\}$ , where  $1 \leq T \leq 8$ , and  $R_i$  belongs to  $N_x$ ,  $1 \leq i \leq T$ . The ordered set, called the  $T$ -tuple, specifies the allowed directions in which the tracer should move. At any point, the tracer looks in the allowed directions in the order in which they are listed in the  $T$ -tuple and moves to the first untraversed figure point encountered. In this way, successive points are chosen to trace the line pattern using a  $T$ -tuple. As soon as the tracer meets a situation at which there is no untraversed figure point in the allowed directions, it makes use of the next  $T$ -tuple given and traces the pattern. For example, the  $T$ -tuples  $F1 = \{5, 7, 6\}$  and  $D1 = \{3, 5, 4\}$  are sufficient to trace



Eight principal directions.

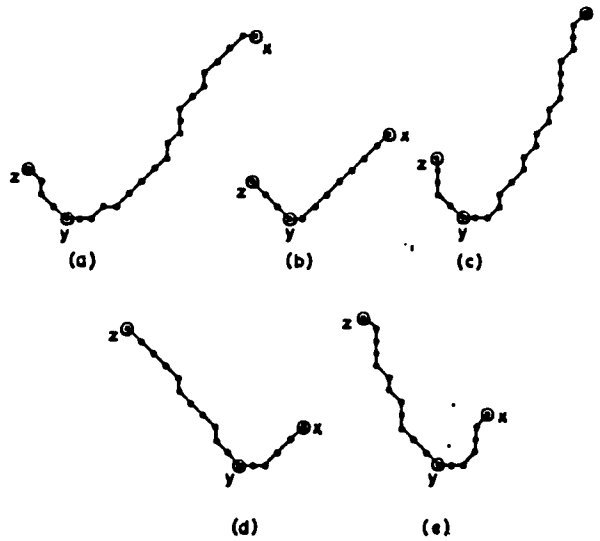


Fig 2.25

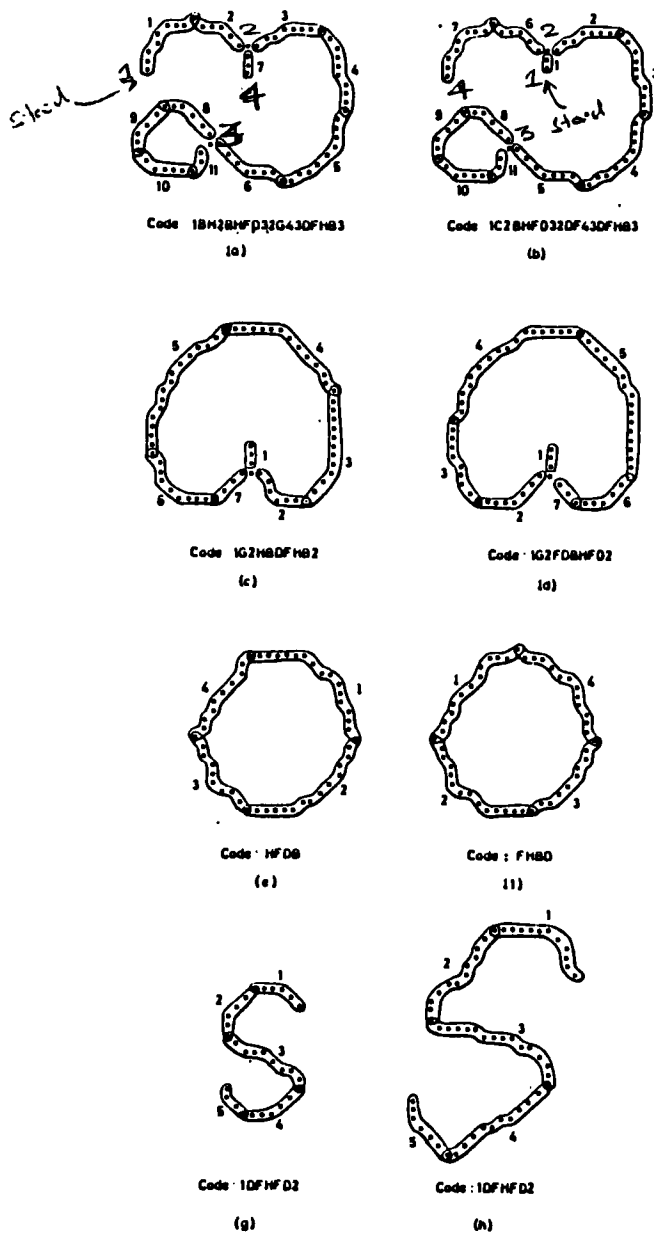
Examples of patterns that can be traced using the T-tuples  $F1$  and  $D1$ . [Raj 72]

$$F1 = \{5, 7, 6\} \text{ and } D1 = \{3, 5, 4\}$$

and recognize the line patterns in Figure 2.25. The tracer starts at x and continues up to y using the T-tuple F1. At the point y, the tracer does not find a figure point in the allowed directions specified in F1. Hence the tracer gets the next T-tuple D1 and traces the pattern. At point z, tracing is stopped.

This method is not affected by variations in size, stretching or squeezing of the line patterns in the horizontal or vertical direction, and is impervious to rotation to a certain extent. However, this method has to be aided by other facts like the relative lengths of line segments traced by different T-tuples, the region of operation of the tracer while using a T-tuple, the relative positions of end points and junction points, etc. Also, the ordering of the directions in a T-tuple should be done with great care. Figure 2.26 shows a build primitive, the T-tuples used, and a flow diagram to direct the tracer.

For recognition of the basic letters, an 'on-the-curve-tracing' procedure was used [Raj 72]. This method involved coding the character in the form of a string of T-tuples by tracing the character along points in it. An appropriate set of T-tuples was chosen to trace the pattern from a set of selected points, like an end point, junction point, top left point, bottom left point, top right point etc. The code represented by the sequence of T-tuples was very compact. (Figure 2.27). A dictionary of codes for the basic letters was prepared including information like the



Examples illustrating the on-the-curve coding procedure.

Figure 2.26 On-The-curve coding procedure. [Raj 72]

length of the curve segments, etc. The code of the test letter was compared with the codes in the dictionary, and a 90% recognition rate was obtained.

It can be seen from the description given in section 2.4 that Telugu is a complicated but structured script which has many curvilinear components. Recognition of similar Indian scripts by [Sin 79], [Chi 80], [Cha 83], and recognition of Telugu characters by Rajasekharan and Deeksnatulu [raj 77], suggests that syntactic or structural methods will be the most appropriate method for Telugu character recognition. With this view in mind, syntactic methods and pattern languages in particular were examined for this study. It was observed that a pattern description language, similar to PLANG, suggested by [Sin 83] would be required. However, such an implementation would be complex and time consuming. Thus this thesis was directed towards statistical methods. The details of the relevant theories and algorithms employed in this thesis are described in Chapter 3.

## CHAPTER 3

### RELEVANT THEORIES AND CONCEPTS

#### 3.1 Relevant Theories And Concepts

In this thesis, the decision theoretic or classificatory approach was used for recognition of Telugu characters. Figure 3.1 describes the functions of the recognition system, which basically consists of two parts, analysis and classification. Analysis consists of feature selection using projection profiles and cross-sections, and adaptive learning to train the system. Classification consists of feature extraction and discrimination using a similarity function.

This chapter contains a detailed description of the various methods of feature extraction, adaptive learning and classification functions used in this study.

#### 3.2 FEATURE SELECTION AND EXTRACTION

The following methods of feature selection were used in this study:

- 1: Application of FFT to the projection profiles
- 2: Cross-sections or column and row runs
- 3: Application of FFT to the cross-sections

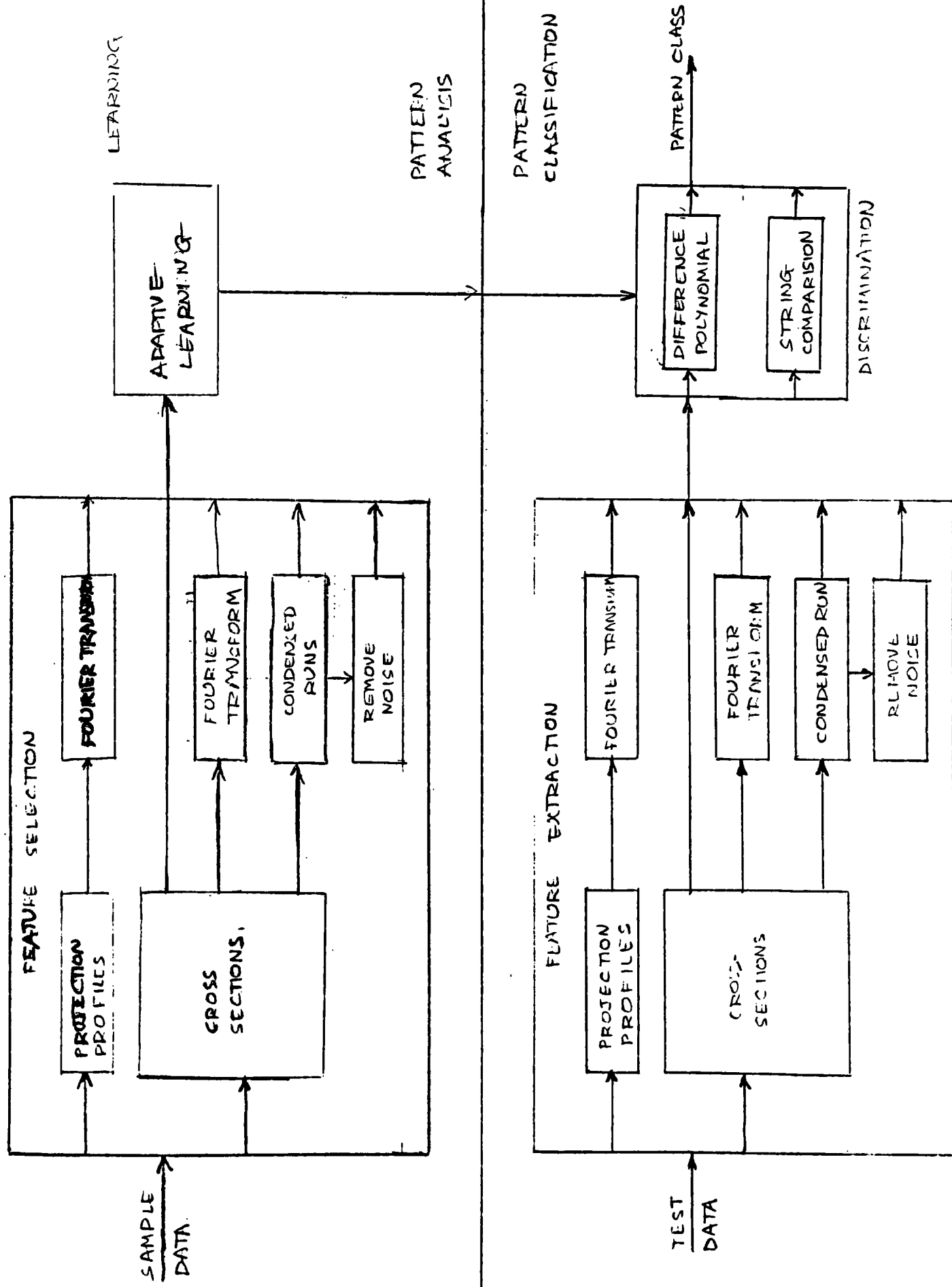


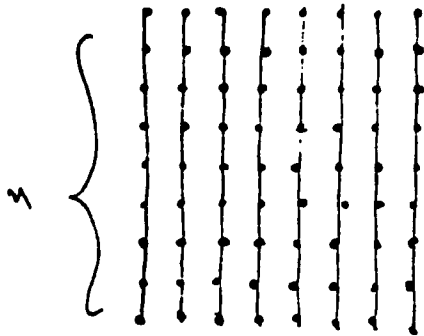
FIG 3.1

FIGURE 3.2

## 2. ALGORITHMS

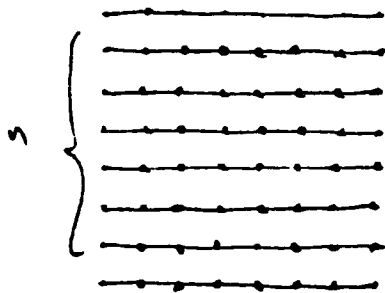
### FEATURE EXTRACTION - METHOD I FOURIER DESCRIPTORS OF PROJECTION PROFILES.

#### 1. VERTICAL PROJECTION PROFILE



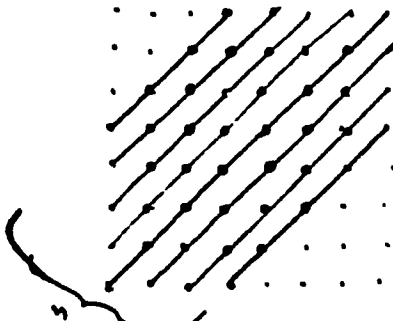
$$V_{SUM}(i) = \sum_{j=1}^N P(i, j)$$

#### 2. HORIZONTAL PROJECTION PROFILE

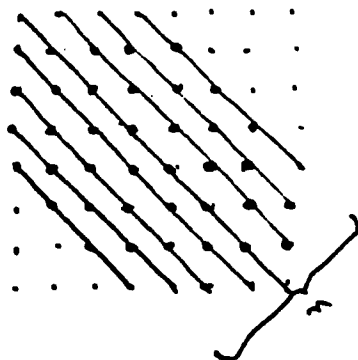


$$H_{SUM}(i) = \sum_{j=1}^N P(j, i)$$

#### 3. ~~RIGHT~~ DIAGONAL PROJECTION PROFILE



#### 4. ~~RIGHT~~ LEFT DIAGONAL PROJECTION PROFILE.



THE CORNERS ARE OMITTED FOR TWO REASONS :- THEY DON'T CONTAIN MUCH USEFUL INFO

- THE OMITTED CORNERS ARE PICKED UP BY OTHER DIAGONAL



4: Condensed cross-sections

5: Condensed cross-sections after noise removal

The same methods were applied for feature extraction also.

### 3.2.1 Projection Profiles -

The projection profiles of an image can be obtained by summing up the grey values in a given direction. For a binary image (black and white), this simplifies to counting the number of black dots (represented as 1's) in a given row or column. Peaks in projection profiles can indicate the locations of major parts of the object. Projection profiles can be generated in different orientations namely rows, column, left diagonal and right diagonal. For Chinese characters, Devanagari script and Tamil characters, which have more horizontal and vertical strokes, the horizontal and vertical projections are more useful. It is not clear, however, how useful these will be for Telugu characters which consist of a lot of curves. It appears that the diagonal projections could be more helpful in this case as they would contain information of the curvilinear components.

Figure 3.2 explains diagrammatically, the methods of obtaining the projection profiles. For a given  $N \times N$  matrix, there will be  $N$  vertical,  $N$  horizontal and  $2N$  diagonal projections. In case of diagonal projections, however, only the center  $N$  lines were considered, i.e., the corner triangles with sides  $N/2 \times n/2$  were omitted. In particular,

the top left and bottom right triangles were omitted for right diagonal projection, and the bottom left and top right triangles were omitted for the left diagonal projection. This reduction was done for the following reasons:

1. The size of diagonal projections would be  $N$ , for an  $N \times N$  matrix instead of  $2N$ .
2. Most of the useful information is usually contained at the center of the picture, hence the corners can be omitted.
3. When both projections are considered, the corners omitted by the right diagonal are included in the left diagonal and vice-versa, without loss of information.

The algorithms for obtaining projection profiles in various directions are described below:

For a binary pattern matrix  $PAT(N \times N)$ :

HORIZONTAL:  $HSUM(i) = PAT(i,j), i = 1,N$

VERTICAL:  $VSUM(i) = PAT(j,i), i = 1,N$

LEFT DIAGONAL:  $LEFT(i) = PAT(i,j), k = (j-i)+N/2$

RIGHT DIAGONAL:  $RIGHT(i) = PAT(i,j), k = (j+i)+N/2$

۱

[illegible]

Figure 3.4 shows the projection profile values for character '౯'. As it can be seen, the projections in the horizontal direction are nothing but the total number of 1's in that row. For example, the projection for row 10 is 3 and for row 11 is 6. Vertical projections for columns 18, 19 and 26 are 1, 16 and 12 respectively.

### 3.2.2 Cross-sections Or Runs: -

Just like projections, cross-sections can give an indication of the characteristics of the pattern. For a binary image, where black dots are represented by 1's and white spaces represented by 0's, a cross-section in a given direction is given by the number of distinct sequences or runs of 1's.

The cross-sections are computed using the following algorithm:

For row runs:


```

FOR I = 1 TO KCWS
  ROWRUN(I) = 0
  FOR J = 1 TO COLUMNS
    IF PATTERN(I,J) = 1 THEN
      IF PATTERN(I-1,J) = 0 THEN
        START A NEW RUN
        ROWRUN(I) = ROWRUN(I) + 1

```

The column runs can be obtained in a similar way.

The number of column runs will be equal to the number of columns and the number of row runs will be equal to the number of rows. This basic set of feature measurements will be unique for each character. For the current study, the first and last four run values were omitted as it was observed that they were mostly zero's, obtaining a reduction in number of features from 64 to 56.

The row and column cross section values for character  are shown in Figure 3.4. In row 10, there are two runs of 1's, one starting at column 27 and the other starting at 30. The length of the run is not considered here; only the number of runs is important. In row 11, there is only one run of 1's starting in column 26. Similarly, in column 20 there is a long but single run of 1's, starting at row 18. Column 29 has three runs starting from 11, 26 and 45.

The column run for this character after deleting the leading and trailing 0's is :

```
1112223333333343332222322333222222211
```

and the row run is:

```
212234444444445544443333222222222222111
```

Typically, the row and column run values are represented without leading and trailing 0's.

### 3.2.3 Condensed Cross-sections -

202303

0000

0000

000000000000011122233333333433322232233322222211000000000000  
COLUMN RUN VALUE

The condensed column-run string is formed by retaining only one digit in each run of that digit in the corresponding column run. This can be done by simply scanning the string and replacing a string of successive duplicate entries by a single entry. The condensed row run is obtained in a similar way.

From Figure 3.4, the full column run for character 'అ' is

1112223333333343332222322333222222211

and its condensed run is

12343232321.

This is obtained by replacing the first 111 by 1, 222 by 2, 33333333 by 3, 4 by 4 etc.

Similarly the row run is

212234444444445544443333222222222222111

and its condensed run is

2123454321.

The advantages of using condensed runs are:

- (1). Condensing is a kind of thinning process.
- (2). The condensed strings will be independent of the thickness of the letters and the proportions of the binary image.
- (3). They require less storage space. (reduction from 64 to

about 16).

(4). The condensed run strings can be unique, qualifying them as good features for recognition.

(5). Easy to compute.

(6). A simple string comparison function can be used for classification.

### 3.2.4 Noise Removal In Condensed Strings -

Noise removal algorithms can be applied after the condensed runs are obtained. The method used in this thesis is derived from Siromoney [Sir 73] and involves removing solitary occurrences of any numbers from the condensed runs, since such occurrences are often due to noise generated in digitization. A perfect example is shown in Figure 3.4 in row 19 at column 29. There is a '0' at this position and the resulting gap is generated by noise. Because of this, now there are two runs of 1's, 11 and 1, though it should have a single run of 1111. This can create variations in cross-sections and lead to rejection by a recognition algorithm. This is a case of an isolated occurrence of a run. It can be seen that if the gap is filled, it will be consistent with the subsequent runs, but it is very difficult to know whether filling the gap would be useful or would create more noise. It was chosen in this thesis to eliminate such occurrences as it was very simple to implement. Another example is at row 27 and column 18. There is a single 1, (or a black dot) in comparison to its



immediate neighbor column 19, which has a run of length 16. This is obviously generated by noise and should be eliminated. This method of eliminating isolated occurrences of runs usually eliminates spurious features being considered. However, there is a possibility that important information could be lost.

Noise removal can be implemented with the help of a count field associated with each value in the condensed cross-section. In the first step, the condensed cross-section is obtained along with the the number of times a numeral occurs in the full run. In the second step, a new string is formed by deleting all the elements from the old string with a count of 1. As a last step, the list is made unique by merging successive elements of the same value, and adding up their counts.

From Figure 3.4, the full column run for character '౯' is

1112223333333343332222322333222222211.

Its condensed run is

12343232321

and its count array is

33813412382.

This is obtained by replacing 111 by 1, and setting the count to 3, replacing 222 by 2 and setting the count to 3,

replacing 33333333 by 3 and setting count to 8, etc. It can be observed that there are two isolated occurrences for 4 and 3 with a count value of 1. After removing these two numerals, the run values are 123322321 with associated counts 338342382. Now, successive runs of 33 can be merged to a single 3, and runs of 22 to a single 2 by adding their associated counts. The final runs values are: 1232321 with counts 3386382, where '8' is 11.

### 3.3 Magnitude Spectra Using Fourier Analysis

Numerical properties of projections, such as their one-dimensional moments (Fourier coefficients), can be useful as object descriptors. Fourier analysis is a general purpose problem solving tool that uses a transform technique from space domain to frequency domain. This is of particular significance to character recognition because the magnitude spectra of a character will be in frequency domains and are independent of the position of the characters. That means that the character can be shifted (but not rotated), and the Fourier descriptors still will be essentially the same. This is sometimes referred to as spatial invariance.

The other advantages of using Fourier transforms (FT's) are that they are simple to use and they provide excellent data reduction. Many transforms give  $N$  descriptors for  $N$  input values. Only  $N/2$  values need be considered because of the

symmetry property of the FT. Even further reduction can be obtained by observing the frequency spectrum and eliminating the less significant components [Bri 78], [Son 85], [Per 77], [Sta 78].

The basic definition of the Fourier transform operation is [Ruc 79]

$$F(\omega) = \int_{-\infty}^{+\infty} f(x) e^{-j\omega x} dx.$$

where:  $F(\omega)$  = the frequency transform

$f(x)$  = the function to be transformed

$\omega$  = the frequency variable (e.g., radians/second)

$x$  = the spatial variable (e.g., time in seconds)

$j$  = sqrt of -1

The transform  $F(\omega)$  is in the complex domain and is performed using complex algebra and integration. The absolute value of  $F(\omega)$  is defined as:

$$|F(\omega)|^2 = F^*(\omega)F(\omega)$$

To calculate  $|F(\omega)|^2$ , we can use the interpretation

$$|F(\omega)|^2 = \{ \text{Re } F(\omega) \}^2 + \{ \text{Im } F(\omega) \}^2.$$

where

$$\begin{aligned} \text{Re } F(\omega) &= \int_{-\infty}^{+\infty} f(x) \cos(\omega x) dx \\ \text{Im } F(\omega) &= \int_{-\infty}^{+\infty} f(x) \sin(\omega x) dx \end{aligned}$$

To finally arrange these equations into a structure suitable for computer calculations consider  $f_i = f(x_i)$  to represent the data value from an experiment or an equation

at position  $x_i$ . For simplicity we assume the data points to be equally spaced, that is  $x_{i+1} - x_i = \Delta x$ , a constant.

We can also consider  $f_i$  to exist only over the interval  $x_1$  to  $x_n$ . Outside that interval,  $f_i$  will be defined to be zero. Combining the above considerations into a programmable form, we have:

$$|F(\omega_j)|^2 = \left\{ \sum_{i=1}^N f_i \cos(\omega_j x_i) \Delta x \right\}^2 + \left\{ \sum_{i=1}^N f_i \sin(\omega_j x_i) \Delta x \right\}^2$$

The magnitude spectrum is simply  $\sqrt{|F(\omega_j)|^2}$ .

This equation is the basis for computer calculation of the Fourier transform of the function represented by  $\{f_i\}$ , in this thesis.

In this thesis, the Fourier transforms were used with projection profiles and cross-sections.

### 3.4 Learning Algorithm

A self adaptive learning algorithm was used in this thesis to train the system with different samples of data so that the various characteristics could be absorbed into the system. This algorithm is a simplified version of that used by Uhr and Vossler [Uhr 63] which adjusted the weights associated with each feature and the values of the features themselves. This method is independent of the values of the features themselves which facilitated using the same learning method for features extracted from different

techniques. This met one of the objectives of the thesis, to study various features and see which one best describes the character. In this study, learning was used with features obtained by Fourier transforms of projection profiles and cross-sections.

The learning algorithm is as follows.

Each character has a set of feature measurements, called typical values, stored in a dictionary. The number of these typical values depends on the method of feature extraction. For each character there is a weight associated with each typical value. Initially, all the weight factors are initialized to a constant value (say 60). The choice of the initial weight factor is important because its value should be neither too low nor too high after learning. The following steps are followed for each feature set to adjust the typical values and weights associated with them.

1. Initialize the weight matrix
2. Input one sample character per standard character to the system.
3. For each input character, adjust the weights of the features of the appropriate characters by comparing the sensitivity or ability of the features to differentiate different characters.

For the kth character, then:

FOR J = 1 TO NUMFEATURES

FOR I = 1 TO NUMSTANDARDCHARACTERS

IF  $TYPICALVALUESAMPLE(J) - TYPICALVALUESTD(K,J) < TYPICALVALUESAMPLE(J) - TYPICALVALUESTD(I,J)$  THEN

THE CORRECT STANDARD CHARACTER IS CLOSER TO THE SAMPLE, SO REWARD IT

$WEIGHT(K,J) = WEIGHT(K,J) + OPTIMUM1$

$WEIGHT(I,J) = WEIGHT(I,J) + OPTIMUM2$

ELSE

THE CORRECT STANDARD CHARACTER IS NOT CLOSER TO THE SAMPLE, SO PUNISH IT

$WEIGHT(K,J) = WEIGHT(K,J) - OPTIMUM1$

$WEIGHT(I,J) = WEIGHT(I,J) - OPTIMUM2$

The values of OPTIMUM1 AND OPTIMUM2 are choosen in such a way that  $OPTIMUM1 > 3*OPTIMUM2$ . This selection basically provides a disctinction between important and less important features.

At this point, after each learning phase, the system can be tested and the performance of recognition and the effect of learning can be studied.

4. The typical value is also adjusted to the average value of all the samples so far. A COUNT is maintained, one for each feature set. Its value is set to one at the time the dictionary is prepared. After every learning phase, its value is bumped by 1.

COUNT = COUNT + 1

TYPICALVALUESTD =

(TYPICALVALUESAMPLE + (TYPICALVALUESTD)\*(COUNT-1))  
/ COUNT

5. This learning process can be repeated for all available sets of samples. Ideally, the larger the number of samples, the better the chances of correct recognition.

### 3.5 Classification:

The classification or assignment of the unknown input to a particular pattern class is accomplished by using a difference polynomial as a measure of similarity. This method was based on the polynomial used by Biles [Bil 80] for pattern analysis and involves the following steps:

1. Normalize weights.

normalizedweight = rawweight/averageweight

where  $\text{averageweight} = \text{sum of weights/number of terms}$

## 2. Compute the difference metric

$\text{totdif} = \text{sum of } \{ \text{normalizedweight} * \text{difference} \}$

where  $\text{difference} =$

$|\text{typicalvalued} - \text{typicalvaluetest}|$

When both patterns are identical,  $\text{totdif}$  will be zero, otherwise its value hopefully will be minimal for the correct standard character. Its value can be adjusted to fall in the range of 0 to 1 if the typical values and their differences are normalized. These values also need to be normalized if the range of typical values is high. In either case, the difference metric with minimum value will hopefully be the correct character, the next higher value will be one for the next closest character etc. Also, it is necessary to normalize the weights, as they vary a lot during adaptive learning.

This difference polynomial was used in this thesis with the fourier descriptors of projection profiles and cross-sections.

A simple string comparison function was used with condensed cross-sections and cross-sections after noise removal.

The results of various methods of feature extraction and the effect of adaptive learning described here will be discussed in Chapter 5.



## CHAPTER 4

### IMPLEMENTATION AND TEST PROCEDURES

#### 4.1 Implementation And Test Procedures

The functional block diagram of the telugu character recognition system is shown in Figure 3.1. The system basically consists of three parts: pattern analysis, learning and recognition.

This chapter contains the description of the hardware configuration used and the software packages developed for the study. The details of test procedures are also described.

#### 4.2 Hardware Configuration

This study used two different hardware configurations: an IBM PC for digitizing the characters and a VAX 11/780 for performing the functions of pattern analysis and recognition. The VAX computer system was selected so that complex analysis and recognition studies could be easily performed with large data sets.

A digital video camera interfaced to an IBM PC was used to digitize images of the characters. The digitized images were stored on a floppy disk and transported to a VAX where the pattern analysis and recognition functions were

performed. Information about some of the available digital video camera equipment can be found in [Cia-1 83], [Cia-2 83].

Figure 4.1 shows the details of the hardware configuration. The equipment used consisted of a digital video camera made by SONY and a video monitor interfaced to the IBM PC through an RS232 port. A powerful software package supplied by IMAGELAB was used to digitize the characters. The characters were written on 8.5X11 inch paper, 16 characters per page, and placed on a flat backlighted board. Another light source was projected directly from the top. The picture was purposefully overexposed to get maximum contrast. A little screen on the video camera was used to adjust the paper within the frame of the video monitor screen. Figures 4.3 shows the set of standard characters and the sample characters used for this study can be found in Figures 4.3 - 4.11. These samples of handprinted characters were collected from three different people.

Once the image was digitized and frozen, several other functions were performed. The image was edited to eliminate any apparant noise around the characters. The grey scales then were adjusted to get maximum contrast and to normalize the values to 0 or 256, with '0' representing a white space and '256' representing a black spot. This was necessary in order to eliminate noise generated due to the texture and unevenness of the paper. Next, the image was zoomed to a

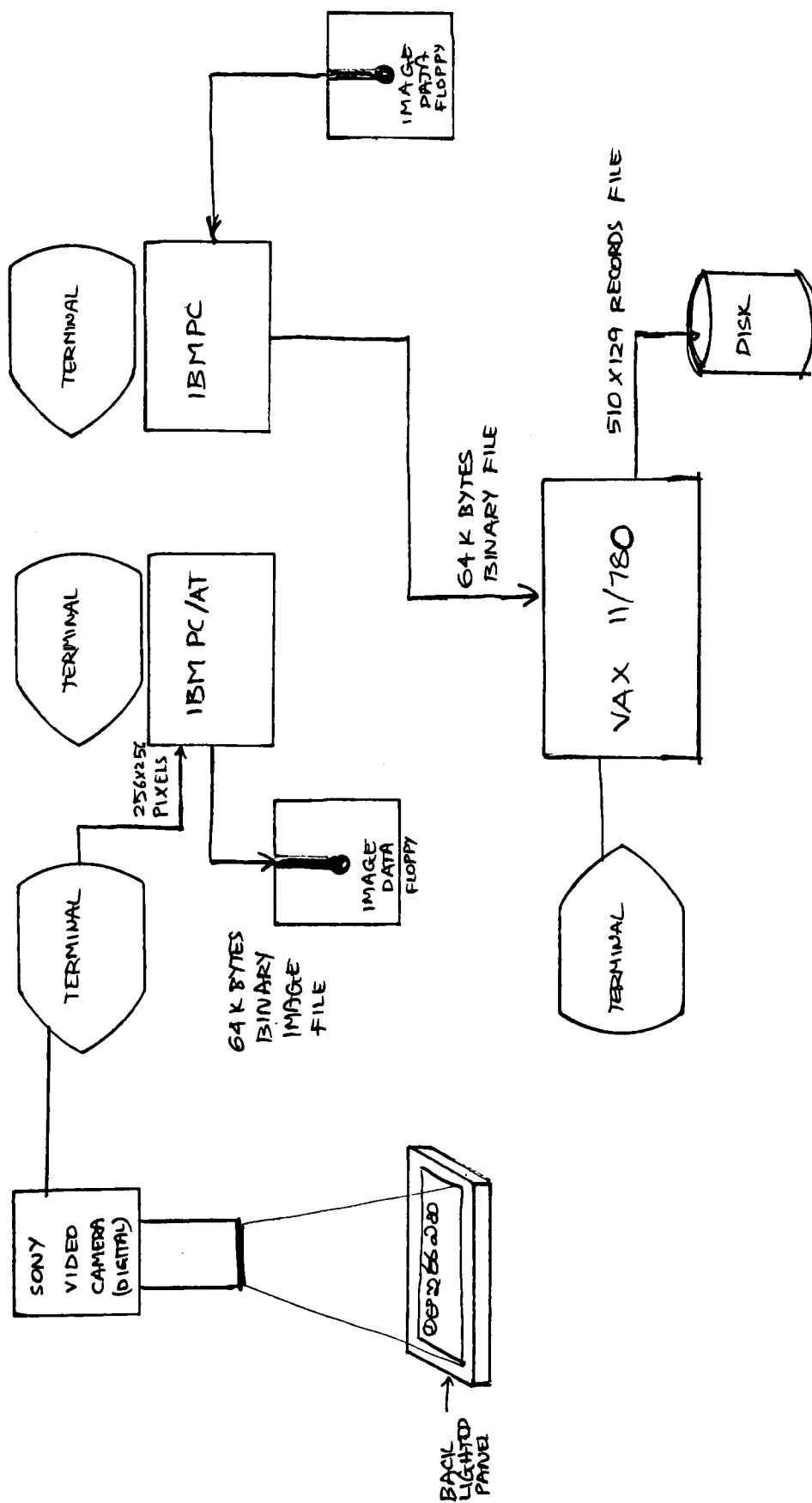


FIGURE 4-1. HARDWARE CONFIGURATION

256X256 pixel size from the original 512X512 size, to get a nice size of 64X64 bytes for each character. Finally, the image was archived on the disk. Each page, containing 16 characters, was stored in a single unformatted file of 64k bytes.

Both standard characters and samples for learning were digitized under the same conditions.

A second IBM PC connected to the VAX was used to upload the image files. All the 64k byte binary unformatted files on the floppy were converted to VMS files of 129 records with a fixed record length of 510 bytes. Format conversion routines were built to extract the binary images of individual characters from these files.

#### 4.3 Software Configuration

The software for analysis, learning and recognition was implemented in FORTRAN 77. FORTRAN is a nice language for scientific, and mathematical applications and was selected as the programming language, because it was easy to implement the Fourier Transform functions. The EQUIVALENCE statement in FORTRAN also came in handy for file format conversions.

The software configuration of the character recognition system is explained in the dataflow diagram in Figure 4.2. The circles represent software functional modules with data flowing in and out of them. Data files are denoted beneath

# TELUGU CHARACTER RECOGNITION

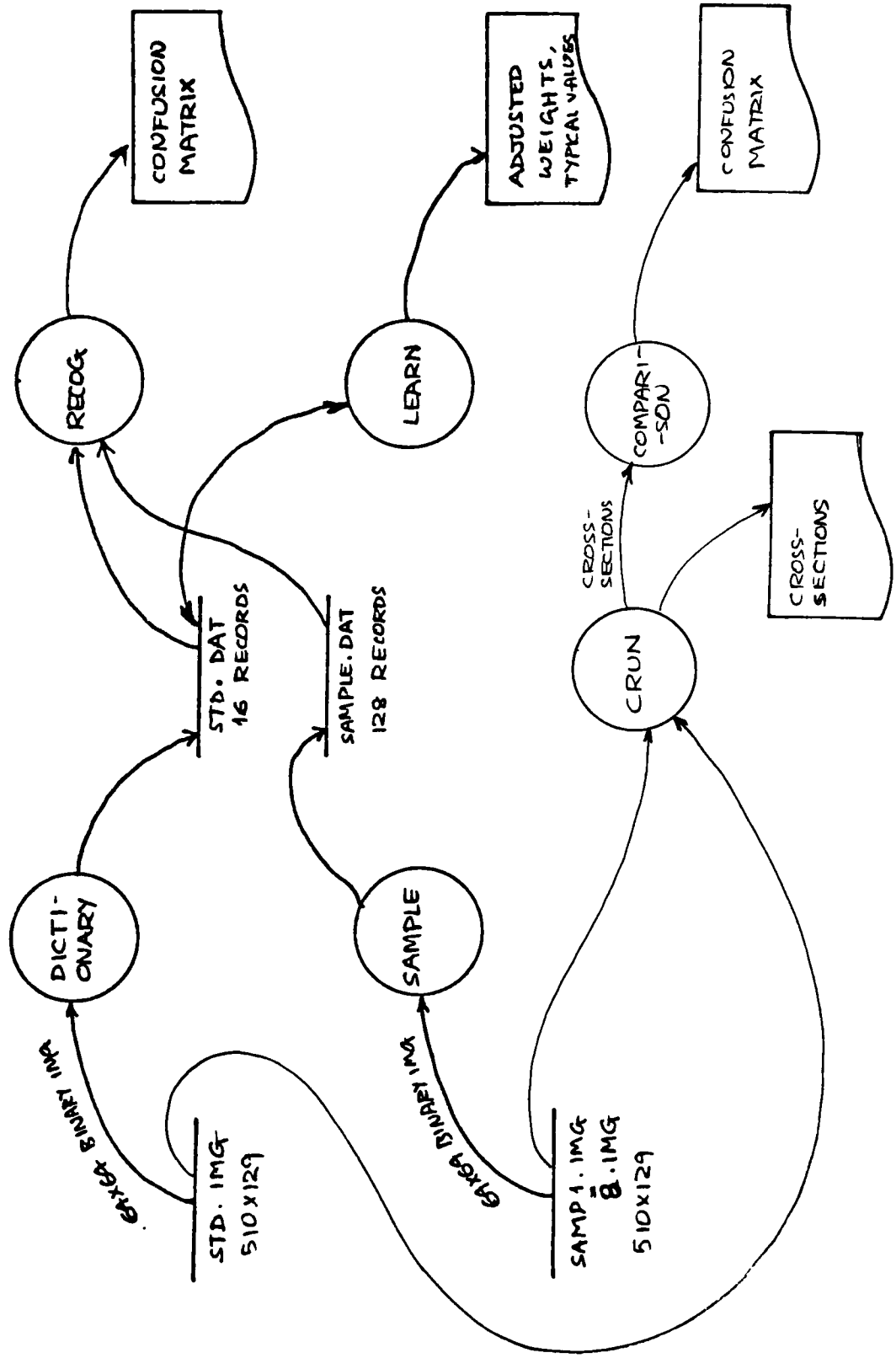


FIG 4-2

horizontal lines. In addition to several test programs, which were developed during the testing phase, the final software basically contained five programs: DICTIONARY, SAMPLE, LEARN, RECOGNIZE AND CRUN. A list of the functions of the programs is given below:

#### DICTIONARY:

Prepares the dictionary of standard characters with features extracted from the binary image file.

Reads the binary image file (129 records of 510 bytes).

For each of the 16 standard characters, computes horizontal, vertical, and diagonal projections and their fourier coefficients, computes column and row cross-sections, and writes one record to per character to std.dat file.

#### SAMPLE:

Prepares sample data file with features extracted from different samples for each standard character.

Reads the binary image file (129 records of 510 bytes).

For each of 128 sample characters, computes horizontal, vertical, and diagonal projections and their fourier coefficients, computes column and row cross-sections, and writes one record to sample.dat file.

#### LEARN:

Uses the self adaptive learning method to adjust the weights of the features of the standard characters to absorb the characteristics of the sample characters.

Reads data from std.dat file and updates it according

to the values read from sample.dat file.

#### RECOGNIZE:

Classifies the test pattern according to the dictionary using a normalized weighted difference polynomial. The correct character will have the minimum value of this polynomial. The values of the polynomial for each of the six features is printed for each standard and for all 128 test characters.

#### CPLN:

Obtains cross-sections of rows and columns, condensed cross-sections and applies noise removal procedure.

#### 4.4 Test Data

A small subset of 16 characters from well over 2000 of the Telugu characters was chosen for this study. The set included a variety of characters which are similar but distinct. They were selected in such a way that some of them should be easy for recognition by projection or films and some by taking cross-sections. Figure 4.5 shows the dictionary characters and Figures 4.3 - 4.11 show handwriting samples of three different people. The standard characters are numbered from 1 to 16 and the samples are numbered from 1 to 128. The characters 'స', 'ష', 'వ', 'ప' are all very similar and are confusing even to human beings. Similarly, the characters 'ల' and 'ళ' differ only at the top right corners but are different vowels. The character 'క' is included as it is totally different from rest of the 15

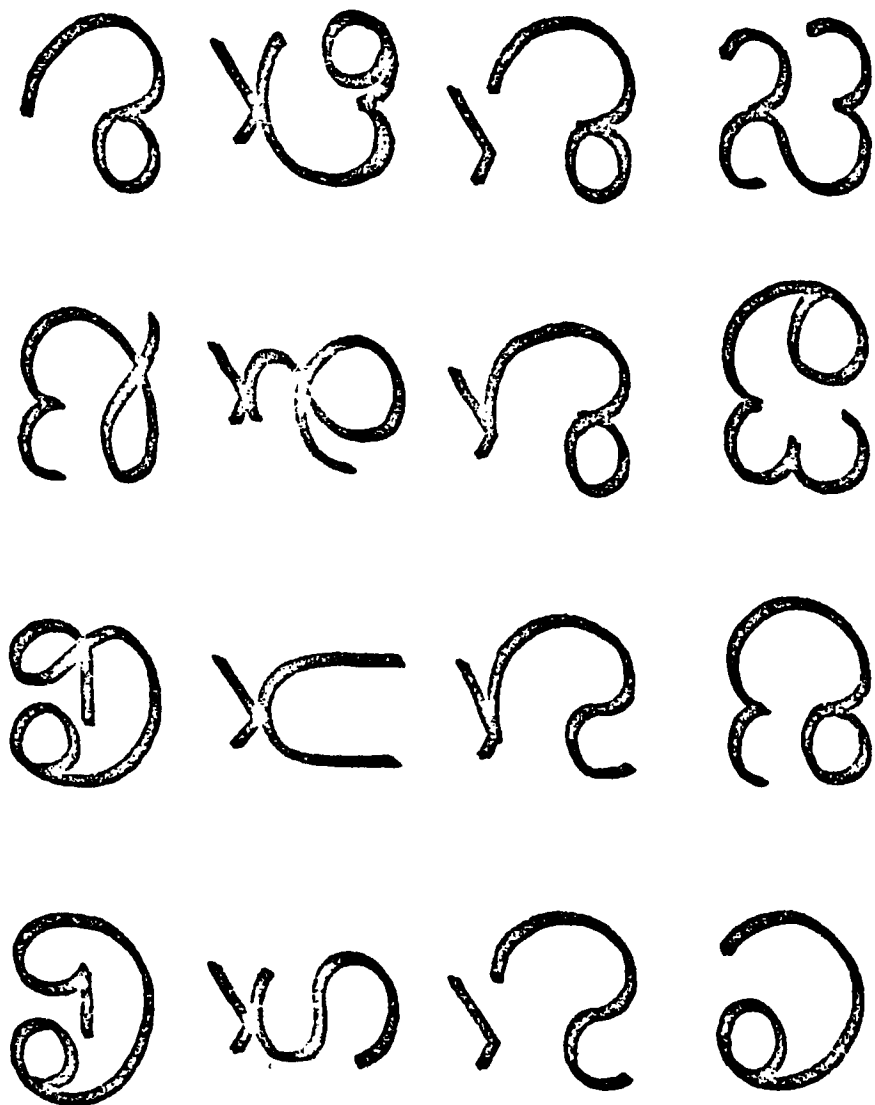


FIG 4.3 STANDARD CHARACTER SET



characters. This data set, thus, is a small but representative sample of most Telugu characters in that there are both easy and hard characters to recognize.

#### 4.5 Test Procedures

A series of tests were performed to analyse the recognition rate of the system with different methods of feature extraction and classification methods.

As discussed above, a dictionary of 16 standard characters was prepared and stored in a data file. A set of 3 samples for each standard character were also digitized. Features were extracted using Fourier analysis of projection profiles and cross sections.

The recognition was tested first without training from samples.

A weighted linear difference polynomial, also called the 'city block' method [Did 76], [Zil 80] was used as the discrimination function. The same 128 samples which were used for training the system were given as the test data.

The next step in testing was to train the system with different handwriting samples. The learning program was run to train the system with one set of samples, one for each standard character, and the classification was tested. The classification program printed the values of the discrimination function. The results of classification were

tabulated in a 'confusion matrix' which shows the systems ability to discriminate different characters. The above steps were repeated 8 times, in order for the system to absorb characteristics of all the samples.

The results of these experiments are described in chapter 5.1

Another series of tests were performed using condensed cross sections as the feature extraction method. In this method, the column and row run values were prepared for both standard and sample characters. The discrimination function was a simple string comparison function to compare the values of the dictionary and test characters. The results are printed in Figure 4.14. The confusion matrix was prepared manually. The same test was repeated after "scrubbing" the data, ie. eliminating solitary occurrences of single runs. These results were also tabulated in a confusion matrix.

## CHAPTER 5

### RESULTS AND DISCUSSION

The objective of this thesis was to study various methods of feature extraction for recognition of cursive Telugu script. This chapter presents the discussion on the results of various experiments. Topics for further research are suggested in Chapter 6.

The binary image of each character was represented by 64X64 bytes. Each of the projection profiles is represented by 64 bytes, and the power spectra of the projection profiles also have the same size. However, using the symmetry property of the Fourier transform, only first 32 values of the Fourier descriptors need be considered. This resulted in a compact representation of the feature set. In addition, from an examination of the magnitude spectra, it could be seen that the lower frequency components are dominant, and that the high frequency components are insignificant. This fact could be used for further data compression of features but was not considered in this study. In all, a total of 128 bytes were required to store all four sets of features for a character.

The values of horizontal, vertical, left diagonal and right diagonal projection profiles and their power spectra are shown in Figure 5.2.

The column and row runs (cross-sections) were represented by 64 bytes each. From an observation of the binary image, it was found that the outermost four lines in both vertical and horizontal directions did not contain any information, so without losing significant data, elements 5 through 50 were selected as the features. This required 112 bytes of space for the two features. In later tests, however, even further data reduction was achieved by taking the power spectrum of the cross-sections and using only the most significant 32 values. Condensed runs required fewer than 32 bytes, and further condensation resulted in even fewer bytes.

The actual values of row and column cross-sections are shown in Figure 5.3 and their fourier descriptors are also shown in Figure 5.3.

The dictionary of standard characters was prepared and stored in a data file. Features of sample characters were extracted in a similar way and the data stored in the sample.dat file. The sample character set used for learning was also used as test input.

The classification program was run with these test samples, and the results are tabulated in what is called a 'confusion matrix'. In this matrix, rows represent the input character and columns represent the recognized character. Thus the diagonal entries correspond to exact matches. These entries would have a maximum value of 0. as

8 samples were used for testing each of the 16 standard characters. A value of  $<8$  indicates incorrect classification, and the mistaken character can be found in the same row.

The recognition rate was calculated as the ratio of the sum of the diagonal elements to the total number of samples. A summary of various test results is presented in Figure 5.1.

### 5.1 Results Of Test 1

The magnitude spectra of horizontal, vertical, left diagonal and right diagonal projection profiles were used as features in this test, and a recognition rate of 64% was obtained. Some of the characters like 'స', 'న', 'ప', 'వ', which are confusing even to humans, were recognized correctly by the system. The confusion matrix for this test is given in Figure. 5.4

The system performed very well after the learning phase, recognizing 96% of the characters. This suggests that the power spectra of projection profiles can be used as a good feature selection method for a cursive script like Telugu. However, in the case of the character 'జ', the system performed worse than its performance before learning. This can be attributed to the fact that the samples are so different that the original characteristics were adversely affected.

FIGURE 5-1.

SUMMARY OF TEST RESULTS

TEST #	FEATURE EXTRACTION METHOD	CLASSIFICATION METHOD	LEARNING	RECOGNITION RATE %
1.	PROTECTION PROFILES FOURIER DESCRIPTORS	LINEAR DIFFERENCE POLYNOMIAL	NO <del>YES</del>	64 %
	"	"	<del>NO</del> YES	96 %
2.	CROSS-SECTIONS	LINEAR DIFFERENCE POLYNOMIAL	NO	20 %
	"	"	YES	20 %
3.	FOURIER DESCRIPTORS CROSS-SECTIONS	"	NO	82 %
	"	"	YES	98 %
4.	CONDENSED <del>COLUMN</del> CROSS-SECTIONS	STRING COMPARISON	NO	32 %
	" ROW	"	NO	8.6 %
5.	CONDENSED COLUMN CROSS-SECTIONS	"	NO	25 %
	" ROW	"	NO	18 %



## 5.2 Results Of Test 2

In this test, the cross-sections of rows and columns were used as features, and a linear difference polynomial was used for classification. As can be seen from the Figure 5.2, the values are very close to each other, making discrimination difficult. Only 20% of the characters were classified correctly, and the performance never seemed to improve even after learning. A 20% recognition rate obviously suggests that the features are not very distinct and unique.

A major contribution to the failure of this method is the spatial variance of the characters, i.e., physical displacement of the characters within a given frame. In the Subsequent tests an attempt was made to eliminate spatial dependencies of the features of the characters.

## 5.3 Results Of Test 3

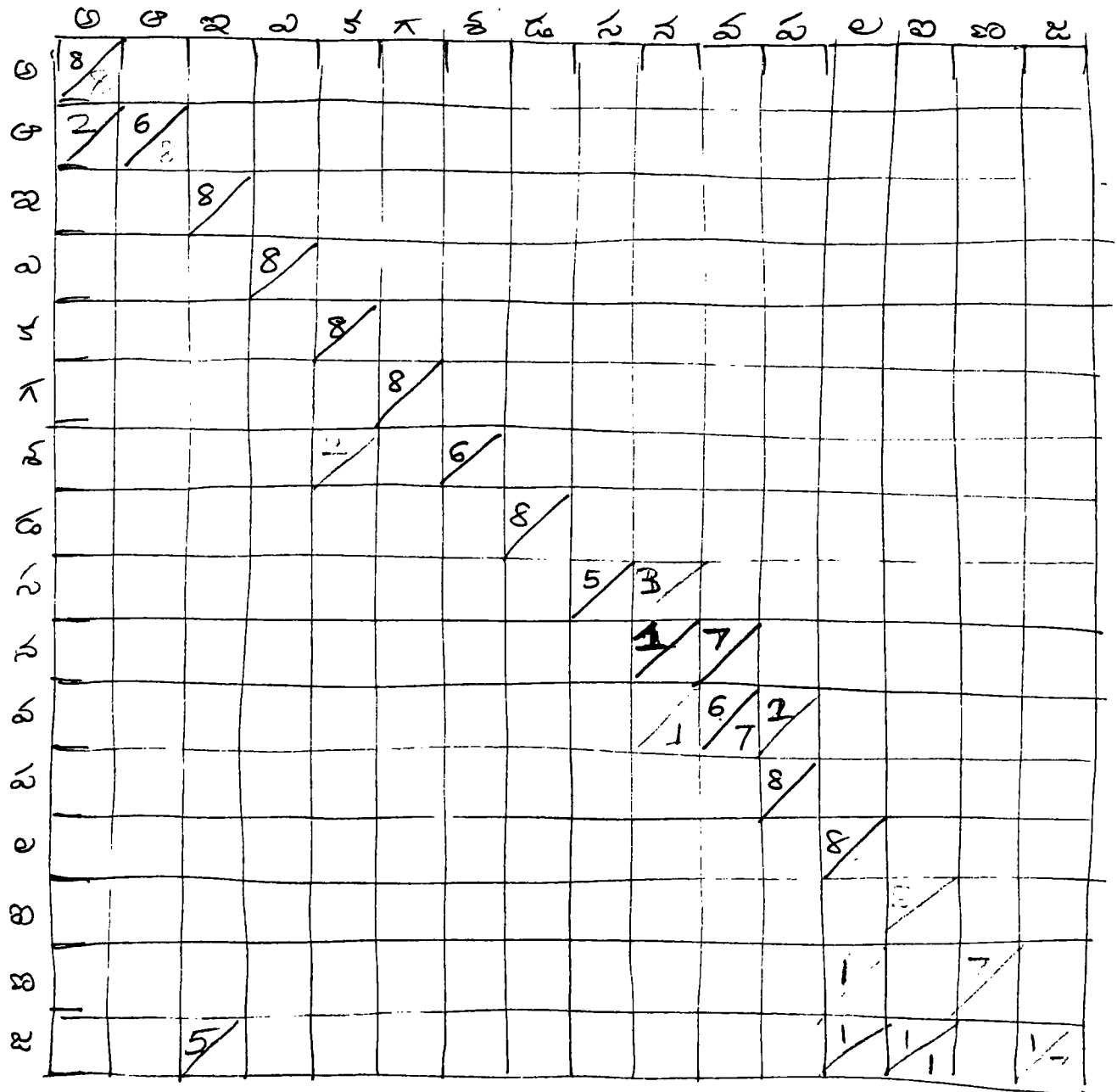
The magintude spectra of the column and row run values were used as features along with a discrimination function method similar to the one used in Test 1. The magintude spectra yielded unique features and an excellent hit rate of 66% before learning and 99% after learning. The Fourier transform eliminated the problem of spatial dependencies, while cross-sections provided unique features, hence the success of this method. See Figure 5.6 for the confusion



# CONFUSION MATRIX

FIG 5.6

RESULTS: MAGNITUDE SPECTRA OF CROSS-SECTIONS



RATE OF RECOGNITION:

BEFORE LEARNING =  $106/128 = 82.8125\%$

AFTER LEARNING =  $126/128 = 98.4375\%$

Fig 5.7

CONFUSION MATRIX  
CROSS-SECTION - COLUMN.

	అ	ఆ	ఇ	ఎ	క	ఁ	ఉ	ఊ	ఋ	ౠ	వ	ప	ల	లి	వి	జ
అ	7/2															
ఆ		2/-														
ఇ			-/3													
ఎ				-/-												
క					1/-											
ఁ						-/3										
ఉ							-/-									
ఊ								-/-								
ఋ									2/-							
ౠ										5/-						
వ											6/8					
ప												5/2				
ల													8/8			
లి														4/5		
వి															1/1	
జ																-/1

RECOGNITION RATE:  $41/128 = 32.03\%$  BEFORE NOISE REDUCTION

$33/128 = 25.78\%$  AFTER NOISE REDUCTION

matrix.

#### 5.4 Results Of Test 4

This was an improvement over test 2, in that the condensed values of column and row runs were used. This method was used successfully by Siromoney [Sir 78] for recognition of printed Tamil characters and was expected to yield better results because the condensed runs have the effect of thinning and characterizing the features. The discrimination method was a simple string comparison. On the average, column runs proved to be more useful than the row runs. The standard characters had unique values, suggesting that this method could be successful. However, the values of the sample runs appear to be affected greatly by the length of the curve, curvature of the curve and relative position of the different segments of the curves in the character. Also, after a close examination of the values, it was found that there were a number of isolated occurrences of single runs, i.e., a feature occurring only once. In general this can be considered to be noise, although it might contain useful features. Improvement can be made by obtaining symbolic runs, removing noise, etc.

Figures 5.7 and 5.8 provide a comparison of how condensed row and column cross-sections scored before and after noise removal.

Fig 5-8

CONFUSION MATRIXCROSS-SECTION ROW

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	-/-															
1		-/1														
2			-/-													
3				2/-												
4					-/6											
5						3/4										
6							1/-									
7								-/-								
8									-/1							
9										1/-						
A											-/1					
B												2/-				
C													1/4			
D														-/2		
E															1/1	
F																-/3

RECOGNITION RATE:  $11/128 = 8.6\%$  BEFORE

$23/128 = 18\%$  AFTER NOISE REDUCTION

### 5.5 Results Of Test 5

Noise was reduced for this test by eliminating isolated occurrences of single runs or cross-section values. This was done at the feature extraction stage, and the method is described in section 3.2.4.

As it was observed in this test, the disadvantage of this noise reduction was that one may lose useful information. This was observed to be the case where the characters had a 'Λ' type of joint between segments.

This method yielded very poor recognition rates of 8%. No learning was attempted in this test, but training the system with different samples might improve the recognition rate.

## CHAPTER 6

### CONCLUSIONS AND FURTHER RESEARCH

#### 6.1 CONCLUSIONS

This chapter describes the advantages of the various character recognition schemes tested in this thesis. Suggestions on expansion and further work in this area will also be presented.

A character recognition system with two different methods of feature selection has been implemented and tested for Telugu character recognition. In the first method, the projection profiles of the pattern, in four different directions (vertical, horizontal, left diagonal and right diagonal) were taken and the Fourier descriptors of these profiles were used as the characteristic features. The system was trained using an adaptive learning algorithm so that characteristics from different handwriting samples could be absorbed. A linear difference polynomial was used as the discrimination function. With the 128 samples of a 16 character set, a 64% recognition rate was obtained before learning which improved to 96% after learning. The diagonal profiles represented the information in the curvilinear components of the characters and in combination with vertical and horizontal profiles, this method yielded very encouraging results.

In the second method, a method of obtaining cross-section information in the rows and columns of the pattern was used for feature extraction. The same linear difference polynomial was used. A very low recognition rate of 20% was obtained. Improvements were tried by obtaining condensed cross-sections and doing noise removal etc., but with little success. Finally, Fourier descriptors of the cross-sections yielded excellent results of 82% recognition rate before learning and 98% after learning. A combination of the Fourier descriptors of column and row cross-sections provided unique representation of the characters under study.

## 6.2 FURTHER STUDY

This study used only 16 from a set of more than 2000 Telugu characters. An obvious extension to this study would be to expand the test character set to include more complex characters and characters of different lengths.

This study yielded excellent results with Fourier descriptors of column and row cross-sections. This can be easily extended to cross-sections of left and right diagonals and should yield promising results since the diagonals would represent features of curvilinear components.

In this study, the use of condensed cross-sections did not yield encouraging results. A more promising method using symbolic cross-sections, used by Siromoney et al [Sir

78], is worth examining. In this method, the normalized lengths of the various segments are maintained.

Another interesting and useful application would be to develop a Telugu script reader. This would require, in addition to recognizing individual characters, a way of identifying a word, with contextual information. This might require a knowledge base of rules to compose words and other structural information.

Telugu characters contain many curvilinear components and are composed of a basic set of primitive symbols using a regular structure. These characteristics suggest that syntactic methods are more suitable for Telugu character recognition [Raj 77]. As discussed above, these methods were considered but rejected for this thesis. Employing syntactic methods [Sin 83] would be another interesting project.

As another project, an electronic facsimile reading system can be developed. This would involve compression of the original image and its later regeneration. Fourier descriptors of projection profiles could give a compact representation of the image. For example, a 64x64 byte image can be reduced to just 32 bytes. Condensed cross-sections can provide even more compact representation (< 32 bytes) and are much faster to generate (no need to do the Fourier analysis). It might be possible to recreate the original image from these compact representations.



The objective of the thesis was achieved by comparing various methods of feature extraction and discrimination. The condensed-cross section method with no adaptive learning scheme did not give encouraging results. However, excellent results were obtained by using magnitude spectra of projection profiles and cross-sections, with cross-sections scoring slightly better than projections.

## BIBLIOGRAPHY

[And 68]. R.H.Anderson, 'Syntax-directed recognition of hand-printed two-dimensional mathematics', Interactive Systems for Experienced Applied Mathematics, ed. by M.Keller and J.Reinfelds (Accademic Press, New York 1968).

[Bar 72]. D.I.Barnea, H.F.Silverman IEEE Transactions on Computers, C-21, 179, 1972.

[Ben 84]. M.Ben-Bassat, L.Zaizendberg 'Contextual template matching; a distance measure for patterns with heirarchically dependent features', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol PAMI-6, NO.2, pp201-11, March 1984.

[Bha 72]. B.K.Bhargava,K.S.Fu, 'Application of Tree system approach to classification of Bubble Chamber photographs', Technical report, TR-EE 72-30, School of Electr. Engg., Purdue University, W.Lafayette, Indiana (Nov 1972).

[Bil 80]. J.A.Biles, 'Adaptive recognition and synthesis of transcribed Jazz Solos by computer', M.S. Thesis, University of Kansas, 1980.

[Bri 76]. E.O.Brigham, "The Fast Fourier Transform", Prentice-Hall, 1976.

[Cha 73]. S.K.Chang, IEEE Trans. Systems, Man and Cybernetics SMC-3, pp.257, 1973.

[Cha 83]. M.Chandrasekharan and R.Chandrasekharan, G.Siromoney, 'Context dependent recognizer for handprinted Tamil characters', Proceedings of 1983 IEEE International Conference on Systems, Man and Cybernetics, Vol 2, pp.1220-4.

[Cha 85]. K.Chandrasekharan, R.Chandrasekharan, G.Siromoney, 'Computer recognition of Tamil, Malayalam and Devanagari characters', Journal Instruments Electronics and Tele com. Engg., (India), Vol 30, NO. 6, pp.150-1546, Nov 1984.

[Chie 75]. Y.P.Chien and K.S.Fu, 'A flexible template matching technique for boundary detection', Technical report TR-EE 75-4, School of Electrical Engg., Purdue University, west Lafayette, Indiana 47907, April 1975.

[Cho 68]. N.Chomsky, M.Halle, The Sound Patterns of English (Harper and Row, New York 1968).

[Chi 80]. P.Chinnuswamy and S.G. Krishnamoorhty, 'Recognition of handprinted Tamil characters', Pattern Recognition, Vol12, pp141-152, 1980.

[Cia-1 83]. Steve Ciarcia, 'Build the Micro D-Cam Solid-state Video Camera, Part 1; The IS32 Optic Ram and Micro D-Cam Hardware', Byte, Sep. 1983, pp.67-86.

[Cia-2 83]. Steve Ciarcia, 'Build the Micro D-Cam Solid-state Video Camera, Part 2; Computer Interfaces and Control Software', Byte, Oct. 1983, pp.269-281.

[Dat 74]. A.K.Datta, 'An experimental procedure for handwritten character recognition', IEEE Trans.Computers, 23, pp.536-545, 1974.

[Dat 85]. A.K.Datta, 'A generalized formal approach for description and analysis of major Indian scripts', Journal Instruments Electronics and Tele com. Engg., (India), Vol 30, NO. 6, pp.155-161, Nov 1984.

[Deu 67]. E.S.Deutsch, 'Computer Simulation of a character-recognition machine, 1 and 2, Post Office Electrical Eng. Journal, 60, pp.39-44,104-109, 1967.

[Did 76]. E.Diday and J.C.Simon, 'Clustering Analysis', pp.47-94, Digital pattern recognition ed.K.S.Fu, 1976

[Dud 77]. S.A.Dudani, K.J.Breeding and R.B.McGhee, 'Aircraft Identification by Moment Invariants', IEEE Transactions on Computers, Vol C-26, pp.39-46, 1977.

[Ede 61]. M.Eden, M.Halle, Proceedings of 4th London symposium Information Theory (Butterworth, London 1961) pp.287-299.

[Fed 68]. J.Feder, Information and Control 13, pp.230, 1968.

[Fre 61]. H.Freeman, IEEE Trans. Elec.Comp. EC-10, pp.260, 1961.

[Fre 62]. H.Freeman, Proc. National Electr. Conf. 16, pp.312, 1962.

[Fu 76a]. K.S.Fu, 'Training in linear classifiers', Syntactic pattern recognition, pp.7-11, ed. K.S.Fu, 1976.

[Fu 76b]. K.S.Fu 'Syntactic pattern recognition and applications' Syntactic pattern recognition, pp.95-135, ed. K.S.Fu, 1976.

[Fu 83]. K.S.Fu, 'Towards unification of syntactic and statistical pattern recognition', Proc. on 3rd Scandinavian Conf. on Image Analysis, July 83, pp.13-22.

[Kan 72]. L.N.Kanal and B.Chandrasekharan, 'On linguistic, statistical and mixed models for pattern recognition', *Frontiers of Pattern Recognition* (S.Watanabe, Ed.), pp.163-192, Accademic Press, New York, 1972.

[Kno 67]. P.J.Knoke and R.J.Wiley, *Proc. IEEE Comp. Conf.*, 1967, p.142.

[Li 83]. H.H.Li and S.C.Chang, 'Projection profiles and Fourier Transform for chinese character recognition', *International Journal Electronics*, Vol 54, No.2, pp.299-309, 1983.

[Luc 83]. D.Lucas, 'Moments technique in picture analysis', *Proceedings of IEEE conference on Computer Vision and Pattern Recognition*, pp.178-87, 1983.

[Mar 84]. S.S.Marwah, S.K.Mullick, R.M.K.Sinha, 'Recognition of Devanagari characters using a heirarchical binary decision tree classifier', *Proceedings of 1984 IEEE Conference on Systems, Man and Cyoernetics*, pp.414-420.

[Moa 75]. B.Moayer and K.S.Fu, 'A syntactic approach to fingerprint pattern recognition', *Pattern Recognition*, Vol 7, 1975, pp.1-23.

[Nag 72]. R.N.Nagel, A.Rosenfield, *Proceedings of IEEE*, 60, p.242, 1972.

[Pav 68]. T.Pavlidis, *Pattern Recognition*, 1, pp.165, 1968.

[Pav-1 72]. T.Pavlidis, *Pattern Recognition*, 4, pp.5, 1972.

[Pav-2 72]. T.Pavlidis, 'Structural pattern recognition', *Frontiers of Pattern recognition*, ed. S.Watanabe, Accademic Press, New York, 1972.

[Per 77]. Eric Persoon, and King-sun Fu, 'Shape discrimination using Fourier descriptors' *IEEE transactions on systems,man and cybernetics*, Vol SMC-7, No.3, March '77, 170-179.

[Pfa 71]. J.L.Ofaltz and A.Rosenfeld, *Proc. First Intern. Joint Conf. Artificial Intelligence*, washington D.C., May 1966, pp.609-619.

[Raj 77]. S.N.S.Rajasekaran and B.L.Deeksnatulu, "Recognition of printed telugu characters", *Computer Graphics and Image Processing* 5, pp335-360 (1977).

[Ray 85]. A.K.Ray and B.Chatterjee, 'Design of a nearest neighbourhood classifier system for Bengali character recognition', *Journal Instruments Electronics and Tele com. Engg.*, (India), Vol 30, NO. 6, pp.226-9, Nov 1984.

[Ros 71]. A.Rosenfield, J.P.Strong, 'A grammar for maps', Software Engineering, Vol 2, ed. J.T.Tou, Accademic Press, New York, 1971.

[Ros 76]. A.Rosenfield and J.S.Weszka, 'Picture Recognition', pp.135-166, Digital pattern recognition, ed.K.S.Fu, 1976.

[Ruc 79]. F.R.Ruckdeschel "Frequency analysis of data using a microcomputer", BYTE, Dec '79, pp.10-35.

[Sha 68]. A.C.Shaw, Rept.SLAC-84, Stanford Linear Accelerator Center, Stanford, California, March 1968.

[Sha 70]. A.C.Shaw, Information and Control 14, p.9, 1969.

[Sha 70]. A.C.Shaw, Journal ACM, 17, p.453, 1970.

[Sin 79]. R.M.K. Sinha, H.N. Mahabala, 'Machine recognition of Devanagari script', IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-9, No.8, Aug 1979.

[Sin 83]. R.M.K. Sinha, 'PLANG - a picture language schema for a class of pictures', Pattern recognition Vol 16, No.4, pp373-383, 1983.

[Sin 84]. R.M.K Sinha, 'A knowledge based script reader', IEEE transactions 1984, pp763-765.

[Sir 76]. G.Siromoney, K.Chandrasekharan and M.Chandrasekharan, 'Computer recognition of printed tamil characters', Pattern Recognition, vol 10,243-247.

[Son 85]. Der-tsyi Song Chuang, 'Adaptive pattern recognition of handwritten chinese characters', M.S. Thesis, K.I.T., 1985.

[Sta 78]. William D. Stanley, "Fast fourier transform on your home computer" BYTE, Dec '78, pp.14-25.

[Tan 84]. G.Y.Tang, P.S.Tzeng,C.C.Hsu, 'A microcomputer system to recognize hanowritten numerals using a syntactic-statistic approach', Seventh Intenational Conference on Pattern Recognition, pp1061-4. Vol 2, 1984.

[Tou 62]. Optical Character Recognition, Spartan books, 1962

[Tou 74]. J.T.Tou and K.C.Gonzalez, Pattern Recognition principles, Addison-Wesley, 1974.

[Uhr 63]. Leonard Uhr and Charles Vossler, 'A pattern recognition program that evaluates and adjusts its own operators', Computers and thought, (E.A. Feigenbaum and J.Feluman, Eds), McGraw-Hill, New York, 1963.

[Vam 73]. T.Vamos and Z.Vassy, 'Industrial pattern recognition experiment - a syntax aide approach', Proc. First Joint Conf.on Pattern Recognition, Oct. 1973.

[Wid 74]. B.Widrow, 'The rubber mask technique', Part 1 and 2, Learning Systems and Intelligent Robots, ed. by K.S.Fu and J.T.Tou, (Plenum Press, New York, 1974).

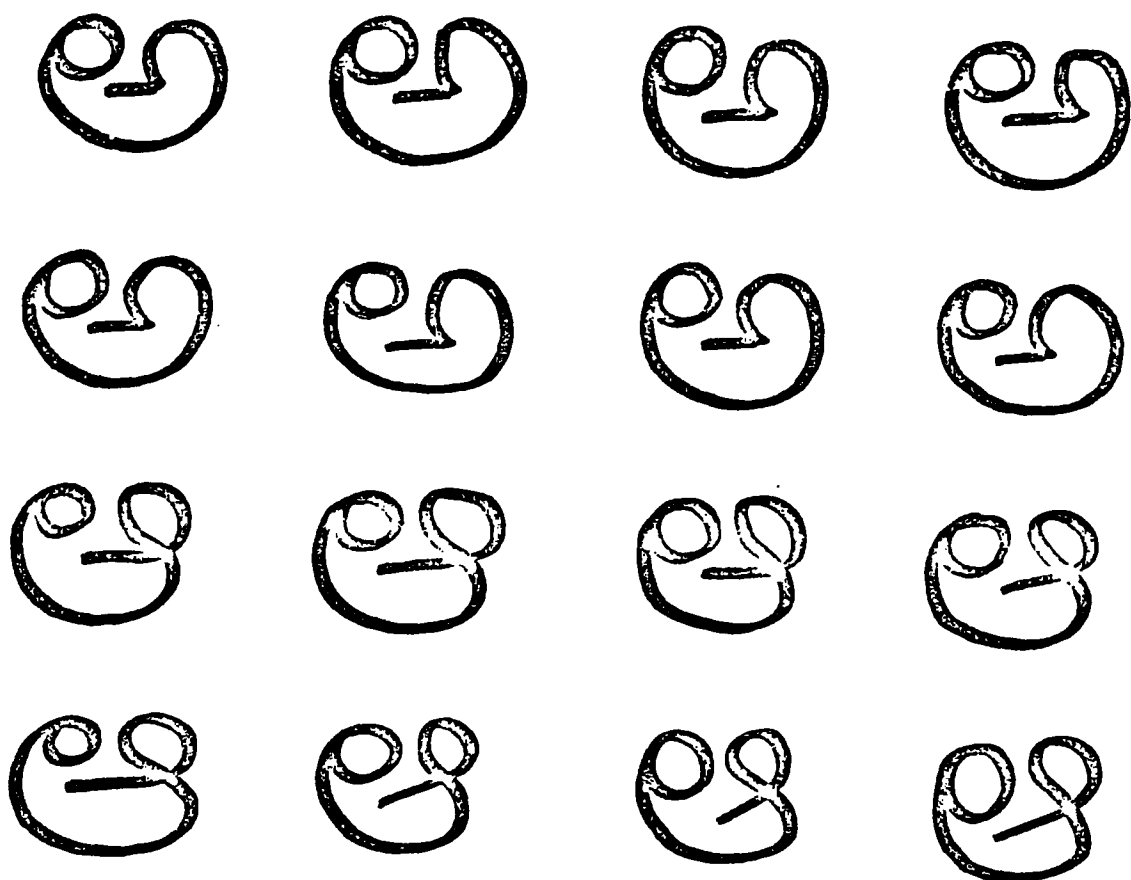


FIG 4.4 SAMPLES FOR 182

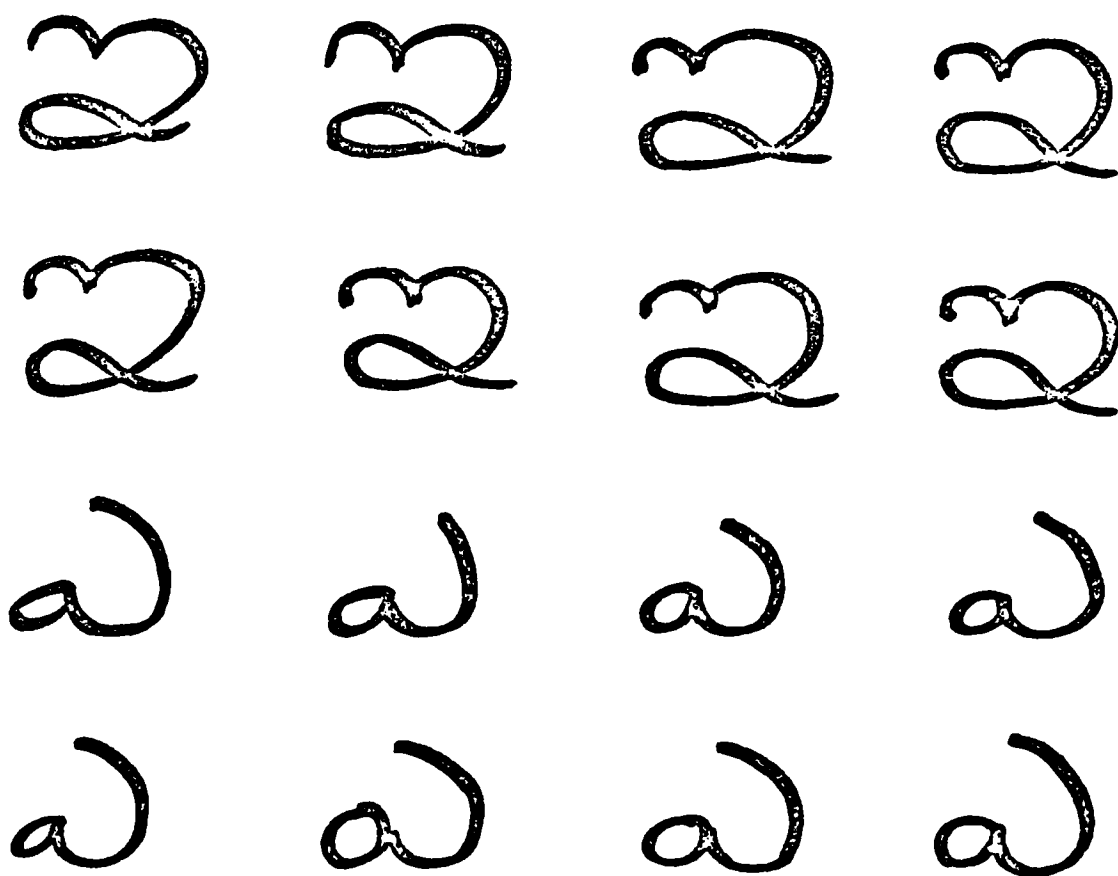


FIG 4.5 SAMPLES FOR 324

u u u u

u u u u

u u u u

u u u u

FIG. 4.6 SAMPLES FOR 586

o o o o

o o o o

o o o o

o o o o

FIG. 4.7 SAMPLES FOR 788



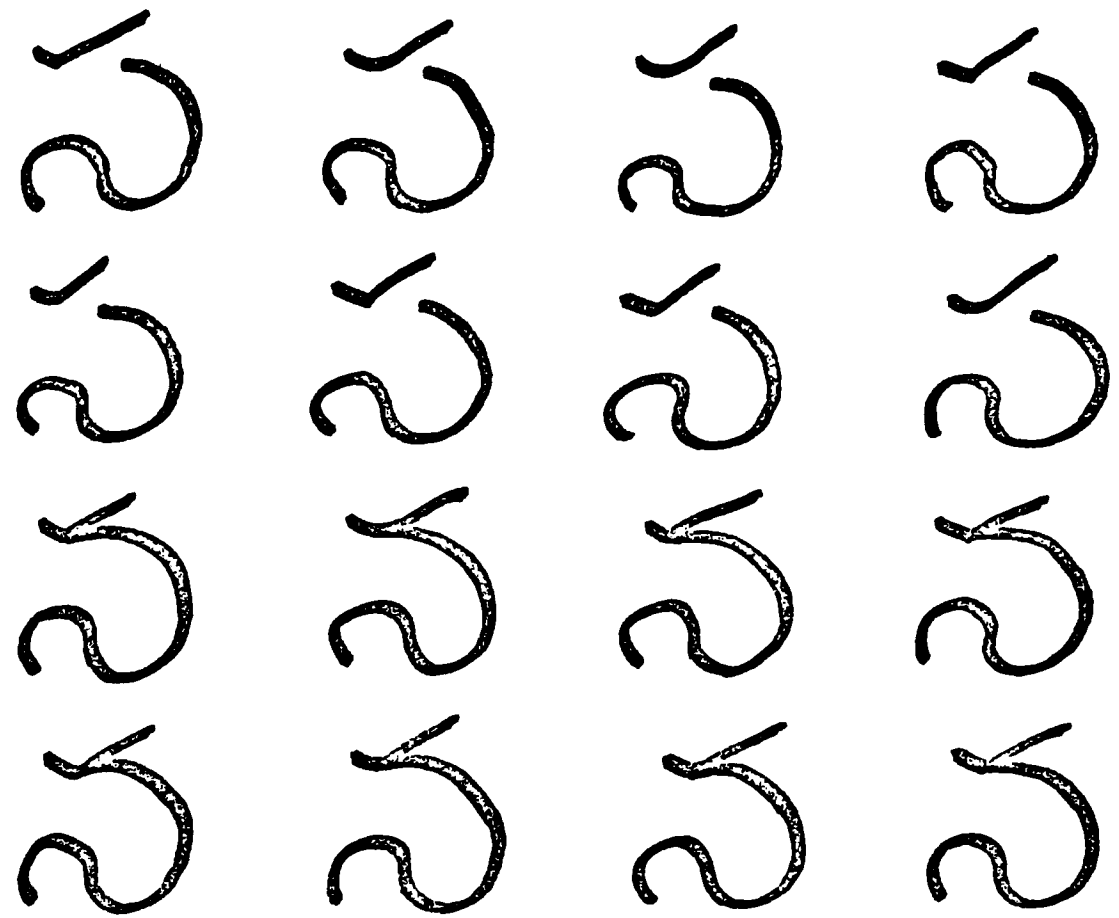


FIG 4-8 SAMPLES FOR 9 and 10

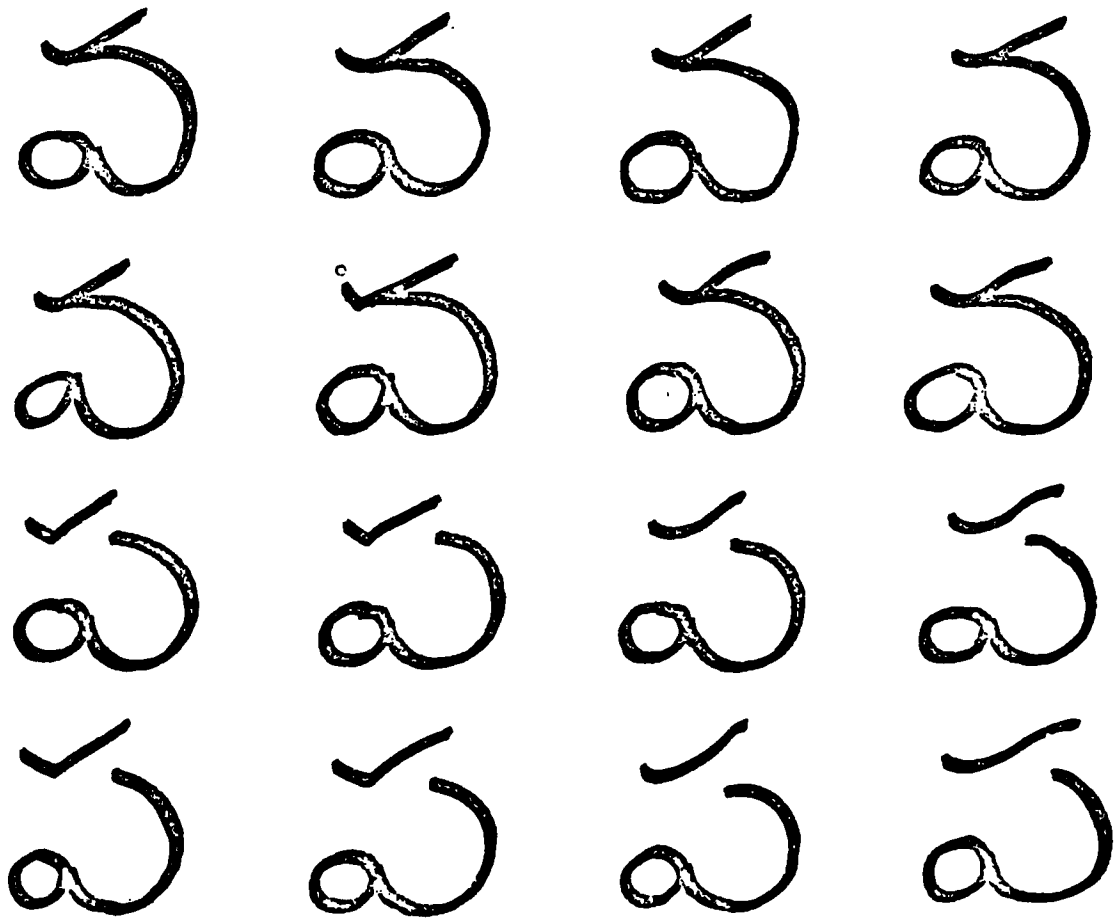


FIG 4-9 SAMPLES FOR 11 AND 12

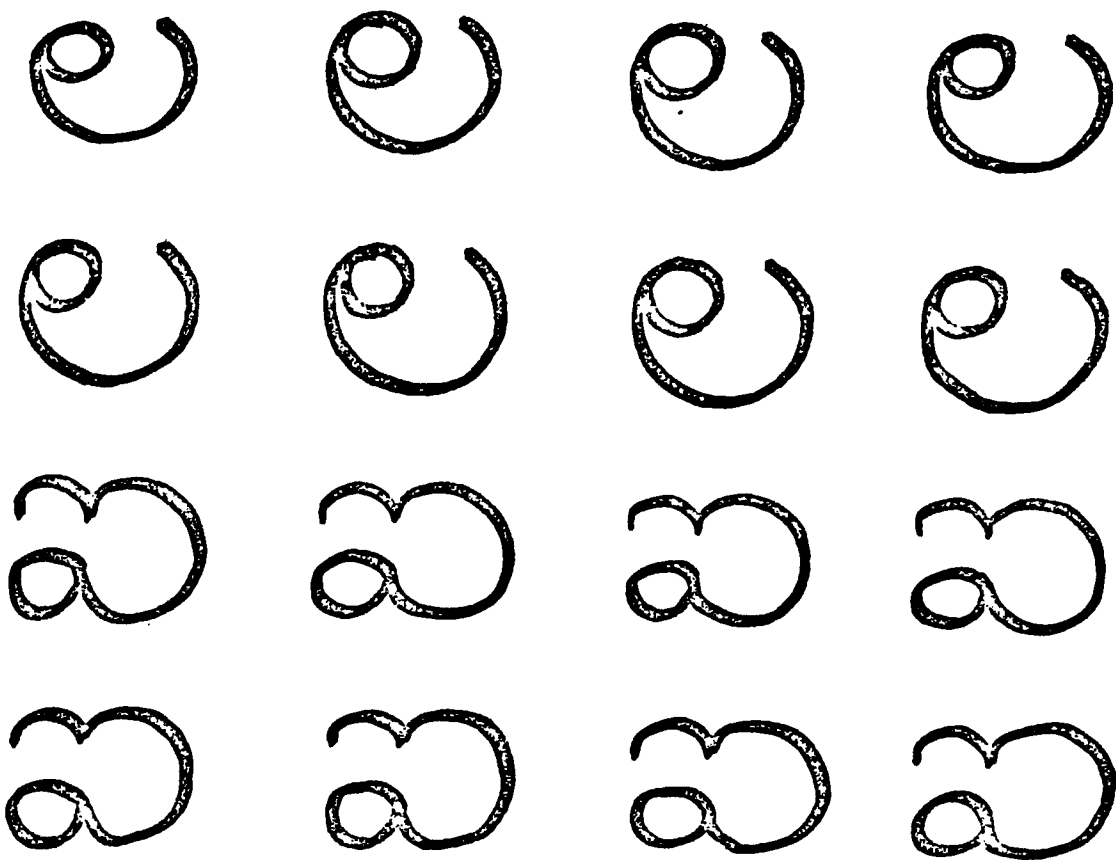


FIG 4-10 SAMPLES FOR 13 & 14

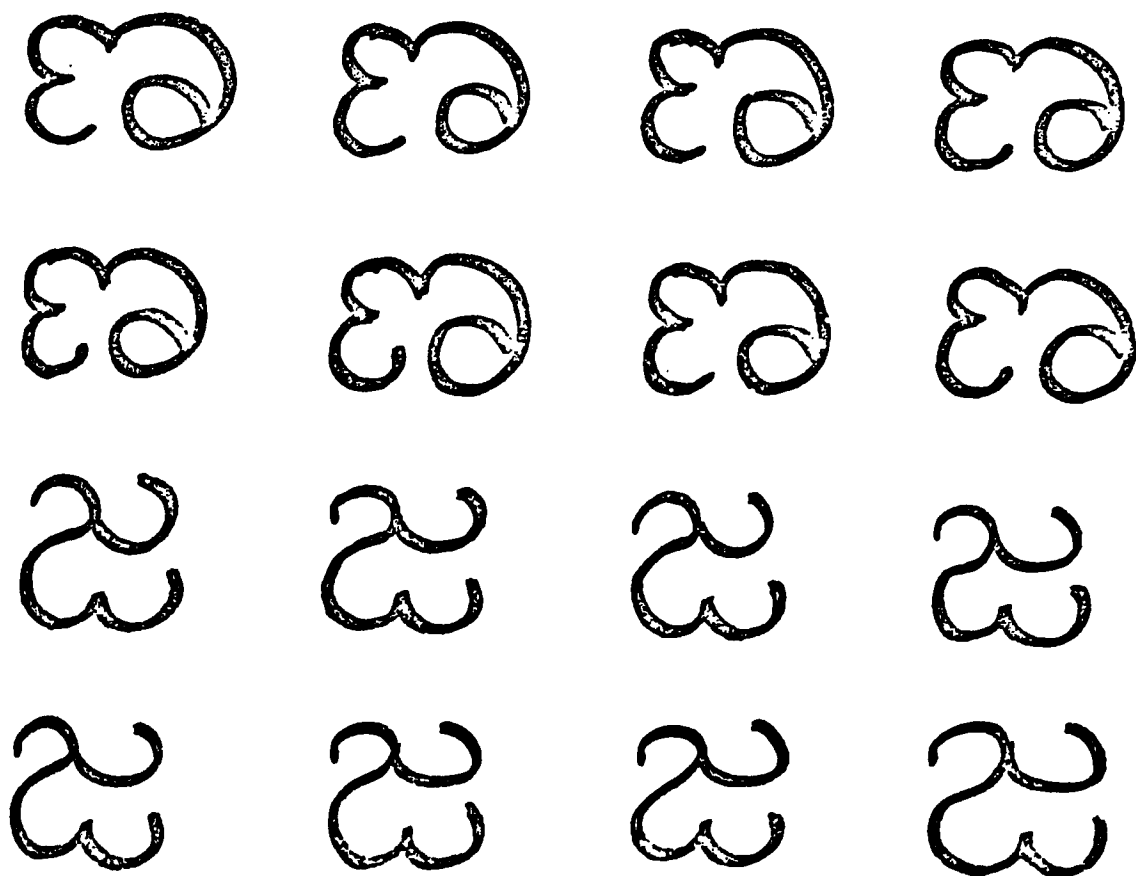


FIG 4-11 SAMPLES FOR 15 & 16

## DATA STRUCTURES

### STD DAT

F.D OF VERTICAL PROJECTION PROFILE	(32)
F.D OF HORIZONTAL "	(32)
F.D OF LEFT DIAGONAL "	(32)
F.D OF RIGHT DIAGONAL "	(32)
COLUMN RUN VALUES	(56)
ROW RUN VALUES	(56)
WEIGHT OF VERTICAL PROJ PROFILE	32
HORIZ	32
LEFT DIAG	32
RIGHT DIAG	32
COLUMN RUN	56
ROW RUN	56
COUNT OF VERTICAL	1
HORIZ	1
LEFT DIAG	1
RIGHT DIAG	1
COLUMN RUN	1
ROW RUN	1

486

### SAMPLE.DAT

Figure 4.12

# OUTPUT OF PATTERN CLASSIFIER

† SUM OF FOUR PROJECTIONS DIFF. POLY.  
 †† SUM OF TWO CROSS-SECTIONS DIFF. POLY.

TELUGU CHARACTER RECOGNITION PROGRAM - REVISION 3

## WEIGHTS OF CHARACTERISTICS

VERITCAL HORIZONTAL

-RUN ROW\_RUN

FILE = 1

CHARACTER NUMBER =

1	195.60	50.38	25.89	36.54	82.79	5.44	1
2	368.84	68.37	129.68	59.48	91.32	17.40	1
3	475.64	192.45	78.69	61.51	142.99	16.64	1
4	520.28	112.49	105.96	92.20	209.63	20.62	1
5	719.79	233.55	161.46	156.62	168.07	38.06	1
6	755.54	251.31	187.48	137.70	179.06	45.44	1
7	656.62	254.59	149.30	107.66	145.07	33.95	1
8	567.45	177.54	103.44	158.26	128.20	18.35	1
9	731.05	221.66	149.78	126.62	233.00	19.90	1
10	751.45	241.89	159.86	116.12	233.58	19.65	1
11	627.89	204.82	104.20	95.37	223.50	9.43	1
12	641.48	190.99	96.63	117.93	235.93	8.02	1
13	545.04	124.06	156.24	113.47	151.28	20.26	1
14	607.40	197.11	111.64	111.08	167.57	9.74	1
15	590.43	154.29	171.70	132.86	131.58	29.64	1
16	533.26	202.50	79.36	99.74	151.68	14.48	1

CHARACTER NUMBER =

1	210.07	16.61	62.01	37.81	93.65	6.88	1
2	329.50	52.06	155.45	63.28	58.73	16.86	1
3	454.62	176.01	114.36	80.76	83.50	13.11	1
4	528.97	106.69	126.81	96.36	199.11	23.55	1
5	707.08	214.97	181.54	144.28	166.30	39.25	1
6	745.67	254.48	193.04	119.65	178.51	49.11	1
7	662.26	242.14	166.18	104.32	149.62	34.81	1
8	540.38	167.74	135.80	135.79	106.55	15.57	1
9	724.58	199.63	165.83	149.12	210.00	22.25	1
10	754.06	212.38	184.41	146.55	210.63	23.21	1
11	600.27	171.14	134.11	98.18	196.50	12.64	1
12	634.61	175.78	111.81	126.11	220.92	11.66	1
13	506.20	111.41	169.24	107.94	117.65	24.35	1
14	550.10	164.53	112.43	110.02	163.12	13.14	1
15	473.59	122.77	139.86	97.04	113.92	27.74	1
16	519.06	168.52	95.34	101.99	153.21	14.70	1

CHARACTER NUMBER =

LEFT-DIAGONAL

RIGHT-DIAGONAL

COLUMN

1	6.31	11.75	1	1
2	12.92	30.32	2	2
3	41.51	58.15	3	3
4	29.65	50.27	4	4
5	50.65	88.71	5	5
6	41.32	86.76	6	6
7	31.78	65.73	7	7
8	31.38	49.74	8	8
9	36.80	56.69	9	9
10	38.81	58.46	10	10
11	37.64	47.07	11	11
12	37.80	45.82	12	12
13	23.89	44.15	13	13
14	37.54	47.29	14	14
15	26.23	55.86	15	15
16	34.74	49.21	16	16

1	5.95	12.83	1	1
2	9.95	26.80	2	2
3	44.39	57.50	3	3
4	31.83	55.38	4	4
5	53.54	92.79	5	5
6	44.14	93.25	6	6
7	31.27	66.07	7	7
8	34.08	49.65	8	8
9	39.17	61.42	9	9
10	40.28	63.49	10	10
11	38.72	51.36	11	11
12	40.95	52.62	12	12
13	24.84	49.19	13	13
14	39.25	52.39	14	14
15	20.85	48.58	15	15
16	34.31	49.01	16	16

### TYPICAL VALUES AFTER LEARNING PHASE - COLUMN RUNS

[illegible]

WEIGHTS OF CONDENSED ROW RUNS AFTER LEARNING

PHASE

## PROGRAM REV 2

NUMBER OF SAMPLES =

WEIGHTS OF CONDENSED COLUMN RUNS AFTER LEARNING

## PHASE

492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492
492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492
492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492
492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492
492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492
492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492
492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492
492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492
492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492	492
432	390	404	258	488	488	488	446	432	460	418	432	446	384	444	418		
306	402	416	430	488	488	488	334	320	390	236	306	306	402	416	416		
268	392	406	434	492	492	492	268	268	212	392	406	268	366	416	392		
172	258	424	362	334	488	404	370	152	222	314	342	208	344	424	372		
324	310	436	412	156	394	394	436	58	422	324	324	324	436	436	436		
224	308	456	448	214	348	418	364	308	448	308	456	224	456	456	456		
92	246	440	456	194	400	432	386	386	162	316	440	316	440	440	440		
280	456	456	476	204	204	464	280	378	476	456	456	280	456	456	456		
240	464	464	464	240	330	254	240	270	464	464	464	464	464	394	464		
178	456	456	472	276	288	268	354	472	472	344	276	456	456	374	456		
472	472	472	484	472	484	282	386	484	484	386	288	472	472	474	388		
476	476	476	484	476	460	302	484	484	484	386	288	476	476	446	406		
468	468	468	376	468	376	176	472	472	472	332	402	468	402	390	402		
456	456	456	488	308	362	166	476	476	420	364	420	456	476	376	420		
452	368	452	492	452	380	156	480	480	424	312	424	452	480	492	424		
356	382	452	492	452	492	156	438	480	438	382	424	452	480	492	424		
426	468	460	366	460	366	156	426	426	426	384	468	330	468	352	468		
304	444	448	352	448	184	416	444	444	402	416	472	430	472	338	472		
416	416	436	180	436	180	416	416	444	388	416	472	434	472	348	472		
324	374	444	194	444	472	444	374	430	318	374	430	448</					

FIG 5.2  
- FDS OF PROJECTION PROFILES - FOR 16 STD CHARACTERS  
- CROSS-SECTIONS -

1/8

FOURIER DESCRIPTIONS OF LEFT DIAGONAL PROJECTION PROFILES															
15.70	11.55	7.34	4.74	4.91	7.33	9.49	10.69	10.74	9.70	4.98	7.67	6.15	7.80	4.21	6.33
10.70	12.71	12.81	12.47	11.20	7.34	6.93	4.51	3.19	4.21	6.15	6.33	6.33	6.33	6.33	6.33
6.12	4.23	1.91	0.64	2.70	4.69	6.14	7.05	7.30	7.09	6.33	5.21	3.98	3.98	3.10	3.95
FOURIER DESCRIPTIONS OF HORIZONTAL PROJECTION PROFILES															
41.97	41.30	39.34	36.19	32.03	27.08	21.61	15.91	10.27	4.98	0.47	3.71	9.25	8.41	7.05	5.30
8.80	7.34	5.55	4.04	3.94	5.04	6.70	8.15	9.50	9.50	9.25	8.41	7.05	5.30	4.65	3.29
0.94	2.70	4.10	5.19	5.71	5.74	5.24	4.34	3.14	1.86	1.30	2.25	3.53	4.65	5.45	5.85
5.84	5.47	4.47	3.53	2.24	0.84	0.74	2.01	3.13	3.98	4.50	4.68	4.52	4.03	3.28	2.33
FOURIER DESCRIPTIONS OF LEFT DIAGONAL PROJECTION PROFILES															
41.47	40.00	35.82	29.52	22.10	14.60	8.45	5.06	5.06	5.43	4.69	3.02	1.61	1.61	2.85	4.69
6.61	4.60	5.01	4.65	3.50	2.30	1.86	2.95	3.77	4.06	3.77	2.42	0.74	1.31	3.12	4.47
5.11	4.60	4.31	2.81	1.54	1.29	1.86	2.13	1.97	1.22	0.96	1.75	2.65	3.24	3.41	3.17
2.63	2.03	1.53	1.32	1.37	1.50	1.59	1.63	1.61	1.55	1.52	1.56	1.61	1.58	1.38	0.96
FOURIER DESCRIPTIONS OF HORIZONTAL PROJECTION PROFILES															
41.51	39.34	36.19	32.03	27.08	21.61	15.91	10.27	4.98	0.47	3.71	9.25	8.41	7.05	5.30	3.29
8.80	7.34	5.55	4.04	3.94	5.04	6.70	8.15	9.50	9.50	9.25	8.41	7.05	5.30	4.65	3.29
0.94	2.70	4.10	5.19	5.71	5.74	5.24	4.34	3.14	1.86	1.30	2.25	3.53	4.65	5.45	5.85
5.84	5.47	4.47	3.53	2.24	0.84	0.74	2.01	3.13	3.98	4.50	4.68	4.52	4.03	3.28	2.33
FOURIER DESCRIPTIONS OF LEFT DIAGONAL PROJECTION PROFILES															
41.47	40.00	35.82	29.52	22.10	14.60	8.45	5.06	5.06	5.43	4.69	3.02	1.61	1.61	2.85	4.69
6.61	4.60	5.01	4.65	3.50	2.30	1.86	2.95	3.77	4.06	3.77	2.42	0.74	1.31	3.12	4.47
5.11	4.60	4.31	2.81	1.54	1.29	1.86	2.13	1.97	1.22	0.96	1.75	2.65	3.24	3.41	3.17
2.63	2.03	1.53	1.32	1.37	1.50	1.59	1.63	1.61	1.55	1.52	1.56	1.61	1.58	1.38	0.96
FOURIER DESCRIPTIONS OF HORIZONTAL PROJECTION PROFILES															
41.51	39.34	36.19	32.03	27.08	21.61	15.91	10.27	4.98	0.47	3.71	9.25	8.41	7.05	5.30	3.29
8.80	7.34	5.55	4.04	3.94	5.04	6.70	8.15	9.50	9.50	9.25	8.41	7.05	5.30	4.65	3.29
0.94	2.70	4.10	5.19	5.71	5.74	5.24	4.34	3.14	1.86	1.30	2.25	3.53	4.65	5.45	5.85
5.84	5.47	4.47	3.53	2.24	0.84	0.74	2.01	3.13	3.98	4.50	4.68	4.52	4.03	3.28	2.33

















FOURIER DESCRIPTORS OF CROSS-SECTIONS  
& PROJECTIONS FOR 16 STD CHARACTERS

INPUT IMAGE FILE = TETRACUTS-1

CHARACTER = 1

FOURIER DESCRIPTORS OF VERTICAL PROJECTION PROFILES															
48.55	47.64	46.63	45.62	44.61	43.60	42.59	41.58	40.57	39.56	38.55	37.54	36.53	35.52	34.51	33.50
4.30	4.27	4.24	4.21	4.18	4.15	4.12	4.09	4.06	4.03	4.00	3.97	3.94	3.91	3.88	3.85
4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97
1.67	2.64	3.61	4.58	5.55	6.52	7.49	8.46	9.43	10.40	11.37	12.34	13.31	14.28	15.25	16.22
48.55	47.65	46.64	45.63	44.62	43.61	42.60	41.59	40.58	39.57	38.56	37.55	36.54	35.53	34.52	33.51
15.01	14.22	13.43	12.64	11.85	11.06	10.27	9.48	8.69	7.90	7.11	6.32	5.53	4.74	3.95	3.16
4.93	4.92	4.91	4.90	4.89	4.88	4.87	4.86	4.85	4.84	4.83	4.82	4.81	4.80	4.79	4.78
5.40	5.94	6.48	7.02	7.56	8.10	8.64	9.18	9.72	10.26	10.80	11.34	11.88	12.42	12.96	13.50
48.55	47.11	45.67	44.23	42.79	41.35	39.91	38.47	37.03	35.59	34.15	32.71	31.27	29.83	28.39	26.95
3.10	2.77	2.44	2.11	1.78	1.45	1.12	0.79	0.46	0.13	-0.20	-0.53	-0.86	-1.19	-1.52	-1.85
4.09	3.71	3.34	2.97	2.60	2.23	1.86	1.49	1.12	0.75	0.38	0.01	-0.36	-0.71	-1.06	-1.41
0.33	1.13	1.76	2.39	3.02	3.65	4.28	4.91	5.54	6.17	6.80	7.43	8.06	8.69	9.32	9.95
47.55	45.53	43.51	41.49	39.47	37.45	35.43	33.41	31.39	29.37	27.35	25.33	23.31	21.29	19.27	17.25
2.83	1.64	4.29	7.64	10.99	14.34	17.69	21.04	24.39	27.74	31.09	34.44	37.79	41.14	44.49	47.84
3.98	2.35	1.37	1.92	2.53	3.14	3.75	4.36	4.97	5.58	6.19	6.80	7.41	8.02	8.63	9.24
2.25	3.20	3.76	4.32	4.88	5.44	6.00	6.56	7.12	7.68	8.24	8.80	9.36	9.92	10.48	11.04
48.87	47.77	46.67	45.57	44.47	43.37	42.27	41.17	40.07	38.97	37.87	36.77	35.67	34.57	33.47	32.37
1.79	1.50	1.21	0.92	0.63	0.34	0.05	-0.24	-0.53	-0.82	-1.11	-1.40	-1.69	-1.98	-2.27	-2.56
0.44	0.37	0.31	0.25	0.19	0.13	0.07	0.01	-0.05	-0.10	-0.15	-0.20	-0.25	-0.30	-0.35	-0.40
0.65	0.59	0.52	0.44	0.36	0.28	0.20	0.12	0.04	-0.04	-0.12	-0.20	-0.28	-0.36	-0.44	-0.52

COLUMN RUN VALUES: 100+1

RUN VALUES: 100+1

CHARACTER = 2

FOURIER DESCRIPTORS OF VERTICAL PROJECTION PROFILES

48.55	47.64	46.63	45.62	44.61	43.60	42.59	41.58	40.57	39.56	38.55	37.54	36.53	35.52	34.51	33.50
4.30	4.27	4.24	4.21	4.18	4.15	4.12	4.09	4.06	4.03	4.00	3.97	3.94	3.91	3.88	3.85
4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97
1.67	2.64	3.61	4.58	5.55	6.52	7.49	8.46	9.43	10.40	11.37	12.34	13.31	14.28	15.25	16.22
48.55	47.65	46.64	45.63	44.62	43.61	42.60	41.59	40.58	39.57	38.56	37.55	36.54	35.53	34.52	33.51
15.01	14.22	13.43	12.64	11.85	11.06	10.27	9.48	8.69	7.90	7.11	6.32	5.53	4.74	3.95	3.16
4.93	4.92	4.91	4.90	4.89	4.88	4.87	4.86	4.85	4.84	4.83	4.82	4.81	4.80	4.79	4.78
5.40	5.94	6.48	7.02	7.56	8.10	8.64	9.18	9.72	10.26	10.80	11.34	11.88	12.42	12.96	13.50
48.55	47.11	45.67	44.23	42.79	41.35	39.91	38.47	37.03	35.59	34.15	32.71	31.27	29.83	28.39	26.95
3.10	2.77	2.44	2.11	1.78	1.45	1.12	0.79	0.46	0.13	-0.20	-0.53	-0.86	-1.19	-1.52	-1.85
4.09	3.71	3.34	2.97	2.60	2.23	1.86	1.49	1.12	0.75	0.38	0.01	-0.36	-0.71	-1.06	-1.41
0.33	1.13	1.76	2.39	3.02	3.65	4.28	4.91	5.54	6.17	6.80	7.43	8.06	8.69	9.32	9.95
47.55	45.53	43.51	41.49	39.47	37.45	35.43	33.41	31.39	29.37	27.35	25.33	23.31	21.29	19.27	17.25
2.83	1.64	4.29	7.64	10.99	14.34	17.69	21.04	24.39	27.74	31.09	34.44	37.79	41.14	44.49	47.84
3.98	2.35	1.37	1.92	2.53	3.14	3.75	4.36	4.97	5.58	6.19	6.80	7.41	8.02	8.63	9.24
2.25	3.20	3.76	4.32	4.88	5.44	6.00	6.56	7.12	7.68	8.24	8.80	9.36	9.92	10.48	11.04
48.87	47.77	46.67	45.57	44.47	43.37	42.27	41.17	40.07	38.97	37.87	36.77	35.67	34.57	33.47	32.37
1.79	1.50	1.21	0.92	0.63	0.34	0.05	-0.24	-0.53	-0.82	-1.11	-1.40	-1.69	-1.98	-2.27	-2.56
0.44	0.37	0.31	0.25	0.19	0.13	0.07	0.01	-0.05	-0.10	-0.15	-0.20	-0.25	-0.30	-0.35	-0.40
0.65	0.59	0.52	0.44	0.36	0.28	0.20	0.12	0.04	-0.04	-0.12	-0.20	-0.28	-0.36	-0.44	-0.52

COLUMN RUN VALUES: 100+1

RUN VALUES: 100+1

CHARACTER = 3

FOURIER DESCRIPTORS OF VERTICAL PROJECTION PROFILES

48.55	47.64	46.63	45.62	44.61	43.60	42.59	41.58	40.57	39.56	38.55	37.54	36.53	35.52	34.51	33.50
4.30	4.27	4.24	4.21	4.18	4.15	4.12	4.09	4.06	4.03	4.00	3.97	3.94	3.91	3.88	3.85
4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97
1.67	2.64	3.61	4.58	5.55	6.52	7.49	8.46	9.43	10.40	11.37	12.34	13.31	14.28	15.25	16.22
48.55	47.65	46.64	45.63	44.62	43.61	42.60	41.59	40.58	39.57	38.56	37.55	36.54	35.53	34.52	33.51
15.01	14.22	13.43	12.64	11.85	11.06	10.27	9.48	8.69	7.90	7.11	6.32	5.53	4.74	3.95	3.16
4.93	4.92	4.91	4.90	4.89	4.88	4.87	4.86	4.85	4.84	4.83	4.82	4.81	4.80	4.79	4.78
5.40	5.94	6.48	7.02	7.56	8.10	8.64	9.18	9.72	10.26	10.80	11.34	11.88	12.42	12.96	13.50
48.55	47.11	45.67	44.23	42.79	41.35	39.91	38.47	37.03	35.59	34.15	32.71	31.27	29.83	28.39	26.95
3.10	2.77	2.44	2.11	1.78	1.45	1.12	0.79	0.46	0.13	-0.20	-0.53	-0.86	-1.19	-1.52	-1.85
4.09	3.71	3.34	2.97	2.60	2.23	1.86	1.49	1.12	0.75	0.38	0.01	-0.36	-0.71	-1.06	-1.41
0.33	1.13	1.76	2.39	3.02	3.65	4.28	4.91	5.54	6.17	6.80	7.43	8.06	8.69	9.32	9.95
47.55	45.53	43.51	41.49	39.47	37.45	35.43	33.41	31.39	29.37	27.35	25.33	23.31	21.29	19.27	17.25
2.83	1.64	4.29	7.64	10.99	14.34	17.69	21.04	24.39	27.74	31.09	34.44	37.79	41.14	44.49	47.84
3.98	2.35	1.37	1.92	2.53	3.14	3.75	4.36	4.97	5.58	6.19	6.80	7.41	8.02	8.63	9.24
2.25	3.20	3.76	4.32	4.88	5.44	6.00	6.56	7.12	7.68	8.24	8.80	9.36	9.92	10.48	11.04
48.87	47.77	46.67	45.57	44.47	43.37	42.27	41.17	40.07	38.97	37.87	36.77	35.67	34.57	33.47	32.37
1.79	1.50	1.21	0.92	0.63	0.34	0.05	-0.24	-0.53	-0.82	-1.11	-1.40	-1.69	-1.98	-2.27	-2.56
0.44	0.37	0.31	0.25	0.19	0.13	0.07	0.01	-0.05	-0.10	-0.15	-0.20	-0.25	-0.30	-0.35	-0.40
0.65	0.59	0.52	0.44	0.36	0.28	0.20	0.12	0.04	-0.04	-0.12	-0.20	-0.28	-0.36	-0.44	-0.52

COLUMN RUN VALUES: 100+1







FOUNDER PROFILES OF RIGHT DIAGONAL PROJECTION PROFILES	27.71	28.40	29.10	29.80	30.50	31.20	31.90	32.60	33.30	34.00	34.70	35.40	36.10	36.80	37.50	38.20	38.90	39.60	40.30	41.00	41.70	42.40	43.10	43.80	44.50	45.20	45.90	46.60	47.30	48.00	48.70	49.40	50.10	50.80	51.50	52.20	52.90	53.60	54.30	55.00	55.70	56.40	57.10	57.80	58.50	59.20	59.90	60.60	61.30	62.00	62.70	63.40	64.10	64.80	65.50	66.20	66.90	67.60	68.30	69.00	69.70	70.40	71.10	71.80	72.50	73.20	73.90	74.60	75.30	76.00	76.70	77.40	78.10	78.80	79.50	80.20	80.90	81.60	82.30	83.00	83.70	84.40	85.10	85.80	86.50	87.20	87.90	88.60	89.30	90.00	90.70	91.40	92.10	92.80	93.50	94.20	94.90	95.60	96.30	97.00	97.70	98.40	99.10	99.80	100.50	101.20	101.90	102.60	103.30	104.00	104.70	105.40	106.10	106.80	107.50	108.20	108.90	109.60	110.30	111.00	111.70	112.40	113.10	113.80	114.50	115.20	115.90	116.60	117.30	118.00	118.70	119.40	120.10	120.80	121.50	122.20	122.90	123.60	124.30	125.00	125.70	126.40	127.10	127.80	128.50	129.20	129.90	130.60	131.30	132.00	132.70	133.40	134.10	134.80	135.50	136.20	136.90	137.60	138.30	139.00	139.70	140.40	141.10	141.80	142.50	143.20	143.90	144.60	145.30	146.00	146.70	147.40	148.10	148.80	149.50	150.20	150.90	151.60	152.30	153.00	153.70	154.40	155.10	155.80	156.50	157.20	157.90	158.60	159.30	160.00	160.70	161.40	162.10	162.80	163.50	164.20	164.90	165.60	166.30	167.00	167.70	168.40	169.10	169.80	170.50	171.20	171.90	172.60	173.30	174.00	174.70	175.40	176.10	176.80	177.50	178.20	178.90	179.60	180.30	181.00	181.70	182.40	183.10	183.80	184.50	185.20	185.90	186.60	187.30	188.00	188.70	189.40	190.10	190.80	191.50	192.20	192.90	193.60	194.30	195.00	195.70	196.40	197.10	197.80	198.50	199.20	199.90	200.60	201.30	202.00	202.70	203.40	204.10	204.80	205.50	206.20	206.90	207.60	208.30	209.00	209.70	210.40	211.10	211.80	212.50	213.20	213.90	214.60	215.30	216.00	216.70	217.40	218.10	218.80	219.50	220.20	220.90	221.60	222.30	223.00	223.70	224.40	225.10	225.80	226.50	227.20	227.90	228.60	229.30	230.00	230.70	231.40	232.10	232.80	233.50	234.20	234.90	235.60	236.30	237.00	237.70	238.40	239.10	239.80	240.50	241.20	241.90	242.60	243.30	244.00	244.70	245.40	246.10	246.80	247.50	248.20	248.90	249.60	250.30	251.00	251.70	252.40	253.10	253.80	254.50	255.20	255.90	256.60	257.30	258.00	258.70	259.40	260.10	260.80	261.50	262.20	262.90	263.60	264.30	265.00	265.70	266.40	267.10	267.80	268.50	269.20	269.90	270.60	271.30	272.00	272.70	273.40	274.10	274.80	275.50	276.20	276.90	277.60	278.30	279.00	279.70	280.40	281.10	281.80	282.50	283.20	283.90	284.60	285.30	286.00	286.70	287.40	288.10	288.80	289.50	290.20	290.90	291.60	292.30	293.00	293.70	294.40	295.10	295.80	296.50	297.20	297.90	298.60	299.30	300.00	300.70	301.40	302.10	302.80	303.50	304.20	304.90	305.60	306.30	307.00	307.70	308.40	309.10	309.80	310.50	311.20	311.90	312.60	313.30	314.00	314.70	315.40	316.10	316.80	317.50	318.20	318.90	319.60	320.30	321.00	321.70	322.40	323.10	323.80	324.50	325.20	325.90	326.60	327.30	328.00	328.70	329.40	330.10	330.80	331.50	332.20	332.90	333.60	334.30	335.00	335.70	336.40	337.10	337.80	338.50	339.20	339.90	340.60	341.30	342.00	342.70	343.40	344.10	344.80	345.50	346.20	346.90	347.60	348.30	349.00	349.70	350.40	351.10	351.80	352.50	353.20	353.90	354.60	355.30	356.00	356.70	357.40	358.10	358.80	359.50	360.20	360.90	361.60	362.30	363.00	363.70	364.40	365.10	365.80	366.50	367.20	367.90	368.60	369.30	370.00	370.70	371.40	372.10	372.80	373.50	374.20	374.90	375.60	376.30	377.00	377.70	378.40	379.10	379.80	380.50	381.20	381.90	382.60	383.30	384.00	384.70	385.40	386.10	386.80	387.50	388.20	388.90	389.60	390.30	391.00	391.70	392.40	393.10	393.80	394.50	395.20	395.90	396.60	397.30	398.00	398.70	399.40	400.10	400.80	401.50	402.20	402.90	403.60	404.30	405.00	405.70	406.40	407.10	407.80	408.50	409.20	409.90	410.60	411.30	412.00	412.70	413.40	414.10	414.80	415.50	416.20	416.90	417.60	418.30	419.00	419.70	420.40	421.10	421.80	422.50	423.20	423.90	424.60	425.30	426.00	426.70	427.40	428.10	428.80	429.50	430.20	430.90	431.60	432.30	433.00	433.70	434.40	435.10	435.80	436.50	437.20	437.90	438.60	439.30	440.00	440.70	441.40	442.10	442.80	443.50	444.20	444.90	445.60	446.30	447.00	447.70	448.40	449.10	449.80	450.50	451.20	451.90	452.60	453.30	454.00	454.70	455.40	456.10	456.80	457.50	458.20	458.90	459.60	460.30	461.00	461.70	462.40	463.10	463.80	464.50	465.20	465.90	466.60	467.30	468.00	468.70	469.40	470.10	470.80	471.50	472.20	472.90	473.60	474.30	475.00	475.70	476.40	477.10	477.80	478.50	479.20	479.90	480.60	481.30	482.00	482.70	483.40	484.10	484.80	485.50	486.20	486.90	487.60	488.30	489.00	489.70	490.40	491.10	491.80	492.50	493.20	493.90	494.60	495.30	496.00	496.70	497.40	498.10	498.80	499.50	500.20	500.90	501.60	502.30	503.00	503.70	504.40	505.10	505.80	506.50	507.20	507.90	508.60	509.30	510.00	510.70	511.40	512.10	512.80	513.50	514.20	514.90	515.60	516.30	517.00	517.70	518.40	519.10	519.80	520.50	521.20	521.90	522.60	523.30	524.00	524.70	525.40	526.10	526.80	527.50	528.20	528.90	529.60	530.30	531.00	531.70	532.40	533.10	533.80	534.50	535.20	535.90	536.60	537.30	538.00	538.70	539.40	540.10	540.80	541.50	542.20	542.90	543.60	544.30	545.00	545.70	546.40	547.10	547.80	548.50	549.20	549.90	550.60	551.30	552.00	552.70	553.40	554.10	554.80	555.50	556.20	556.90	557.60	558.30	559.00	559.70	560.40	561.10	561.80	562.50	563.20	563.90	564.60	565.30	566.00	566.70	567.40	568.10	568.80	569.50	570.20	570.90	571.60	572.30	573.00	573.70	574.40	575.10	575.80	576.50	577.20	577.90	578.60	579.30	580.00	580.70	581.40	582.10	582.80	583.50	584.20	584.90	585.60	586.30	587.00	587.70	588.40	589.10	589.80	590.50	591.20	591.90	592.60	593.30	594.00	594.70	595.40	596.10	596.80	597.50	598.20	598.90	599.60	600.30	601.00	601.70	602.40	603.10	603.80	604.50	605.20	605.90	606.60	607.30	608.00	608.70	609.40	610.10	610.80	611.50	612.20	612.90	613.60	614.30	615.00	615.70	616.40	617.10	617.80	618.50	619.20	619.90	620.60	621.30	622.00	622.70	623.40	624.10	624.80	625.50	626.20	626.90	627.60	628.30	629.00	629.70	630.40	631.10	631.80	632.50	633.20	633.90	634.60	635.30	636.00	636.70	637.40	638.10	638.80	639.50	640.20	640.90	641.60	642.30	643.00	643.70	644.40	645.10	645.80	646.50	647.20	647.90	648.60	649.30	650.00	650.70	651.40	652.10	652.80	653.50	654.20	654.90	655.60	656.30	657.00	657.70	658.40	659.10	659.80	660.50	661.20	661.90	662.60	663.30	664.00	664.70	665.40	666.10	666.80	667.50	668.20	668.90	669.60	670.30	671.00	671.70	672.40	673.10	673.80	674.50	675.20	675.90	676.60	677.30	678.00	678.70	679.40	680.10	680.80	681.50	682.20	682.90	683.60	684.30	685.00	685.70	686.40	687.10	687.80	688.50	689.20	689.90	690.60	691.30	692.00	692.70	693.40	694.10	694.80	695.50	696.20	696.90	697.60	698.30	699.00	699.70	700.40	701.10	701.80	702.50	703.20	703.90	704.60	705.30	706.00	706.70	707.40	708.10	708.80	709.50	710.20	710.90	711.60	712.30	713.00	713.70	714.40	715.10	715.80	716.50	717.20	717.90	718.60	719.30	720.00	720.70	721.40	722.10	722.80	723.50	724.20	724.90	725.60	726.30	727.00	727.70	728.40	729.10	729.80	730.50	731.20	731.90	732.60	733.30	734.00	734.70	735.40	736.10	736.80	737.50	738.20	738.90	739.60	740.30	741.00	741.70	742.40	743.10	743.80	744.50	745.20	745.90	746.60	747.30	748.00	748.70	749.40	750.10	750.80	751.50	752.20	752.90	753.60	754.30	755.00	755.70	756.40	757.10	757.80	758.50	759.20	759.90	760.60	761.30	762.00	762.70	763.40	764.10	764.80	765.50	766.20	766.90	767.60	768.30	769.00	769.70	770.40	771.10	771.80	772.50	773.20	773.90	774.60	775.30	776.00	776.70	777.40	778.10	778.80	779.50	780.20	780.90	781.60	782.30	783.00	783.70	784.40	785.10	785.80	786.50	787.20	787.90	788.60	789.30	790.00	790.70	791.40	792.10	792.80	793.50	794.20	794.90	795.60	796.30	797.00	797.70	798.40	799.10	799.80	800.50	801.20	801.90	802.60	803.30	804.00	804.70	805.40	806.10	806.80	807.50	808.20	808.90	809.60	810.30	811.00	811.70	812.40	813.10	813.80	814.50	815.20	815.90	816.60	817.30	818.00	818.70</
--	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	----------

0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45	1.50	1.55	1.60	1.65	1.70	1.75	1.80	1.85	1.90	1.95	2.00	2.05	2.10	2.15	2.20	2.25	2.30	2.35	2.40	2.45	2.50	2.55	2.60	2.65	2.70	2.75	2.80	2.85	2.90	2.95	3.00	3.05	3.10	3.15	3.20	3.25	3.30	3.35	3.40	3.45	3.50	3.55	3.60	3.65	3.70	3.75	3.80	3.85	3.90	3.95	4.00	4.05	4.10	4.15	4.20	4.25	4.30	4.35	4.40	4.45	4.50	4.55	4.60	4.65	4.70	4.75	4.80	4.85	4.90	4.95	5.00	5.05	5.10	5.15	5.20	5.25	5.30	5.35	5.40	5.45	5.50	5.55	5.60	5.65	5.70	5.75	5.80	5.85	5.90	5.95	6.00	6.05	6.10	6.15	6.20	6.25	6.30	6.35	6.40	6.45	6.50	6.55	6.60	6.65	6.70	6.75	6.80	6.85	6.90	6.95	7.00	7.05	7.10	7.15	7.20	7.25	7.30	7.35	7.40	7.45	7.50	7.55	7.60	7.65	7.70	7.75	7.80	7.85	7.90	7.95	8.00	8.05	8.10	8.15	8.20	8.25	8.30	8.35	8.40	8.45	8.50	8.55	8.60	8.65	8.70	8.75	8.80	8.85	8.90	8.95	9.00	9.05	9.10	9.15	9.20	9.25	9.30	9.35	9.40	9.45	9.50	9.55	9.60	9.65	9.70	9.75	9.80	9.85	9.90	9.95	10.00	10.05	10.10	10.15	10.20	10.25	10.30	10.35	10.40	10.45	10.50	10.55	10.60	10.65	10.70	10.75	10.80	10.85	10.90	10.95	11.00	11.05	11.10	11.15	11.20	11.25	11.30	11.35	11.40	11.45	11.50	11.55	11.60	11.65	11.70	11.75	11.80	11.85	11.90	11.95	12.00	12.05	12.10	12.15	12.20	12.25	12.30	12.35	12.40	12.45	12.50	12.55	12.60	12.65	12.70	12.75	12.80	12.85	12.90	12.95	13.00	13.05	13.10	13.15	13.20	13.25	13.30	13.35	13.40	13.45	13.50	13.55	13.60	13.65	13.70	13.75	13.80	13.85	13.90	13.95	14.00	14.05	14.10	14.15	14.20	14.25	14.30	14.35	14.40	14.45	14.50	14.55	14.60	14.65	14.70	14.75	14.80	14.85	14.90	14.95	15.00	15.05	15.10	15.15	15.20	15.25	15.30	15.35	15.40	15.45	15.50	15.55	15.60	15.65	15.70	15.75	15.80	15.85	15.90	15.95	16.00	16.05	16.10	16.15	16.20	16.25	16.30	16.35	16.40	16.45	16.50	16.55	16.60	16.65	16.70	16.75	16.80	16.85	16.90	16.95	17.00	17.05	17.10	17.15	17.20	17.25	17.30	17.35	17.40	17.45	17.50	17.55	17.60	17.65	17.70	17.75	17.80	17.85	17.90	17.95	18.00	18.05	18.10	18.15	18.20	18.25	18.30	18.35	18.40	18.45	18.50	18.55	18.60	18.65	18.70	18.75	18.80	18.85	18.90	18.95	19.00	19.05	19.10	19.15	19.20	19.25	19.30	19.35	19.40	19.45	19.50	19.55	19.60	19.65	19.70	19.75	19.80	19.85	19.90	19.95	20.00	20.05	20.10	20.15	20.20	20.25	20.30	20.35	20.40	20.45	20.50	20.55	20.60	20.65	20.70	20.75	20.80	20.85	20.90	20.95	21.00	21.05	21.10	21.15	21.20	21.25	21.30	21.35	21.40	21.45	21.50	21.55	21.60	21.65	21.70	21.75	21.80	21.85	21.90	21.95	22.00	22.05	22.10	22.15	22.20	22.25	22.30	22.35	22.40	22.45	22.50	22.55	22.60	22.65	22.70	22.75	22.80	22.85	22.90	22.95	23.00	23.05	23.10	23.15	23.20	23.25	23.30	23.35	23.40	23.45	23.50	23.55	23.60	23.65	23.70	23.75	23.80	23.85	23.90	23.95	24.00	24.05	24.10	24.15	24.20	24.25	24.30	24.35	24.40	24.45	24.50	24.55	24.60	24.65	24.70	24.75	24.80	24.85	24.90	24.95	25.00	25.05	25.10	25.15	25.20	25.25	25.30	25.35	25.40	25.45	25.50	25.55	25.60	25.65	25.70	25.75	25.80	25.85	25.90	25.95	26.00	26.05	26.10	26.15	26.20	26.25	26.30	26.35	26.40	26.45	26.50	26.55	26.60	26.65	26.70	26.75	26.80	26.85	26.90	26.95	27.00	27.05	27.10	27.15	27.20	27.25	27.30	27.35	27.40	27.45	27.50	27.55	27.60	27.65	27.70	27.75	27.80	27.85	27.90	27.95	28.00	28.05	28.10	28.15	28.20	28.25	28.30	28.35	28.40	28.45	28.50	28.55	28.60	28.65	28.70	28.75	28.80	28.85	28.90	28.95	29.00	29.05	29.10	29.15	29.20	29.25	29.30	29.35	29.40	29.45	29.50	29.55	29.60	29.65	29.70	29.75	29.80	29.85	29.90	29.95	30.00	30.05	30.10	30.15	30.20	30.25	30.30	30.35	30.40	30.45	30.50	30.55	30.60	30.65	30.70	30.75	30.80	30.85	30.90	30.95	31.00	31.05	31.10	31.15	31.20	31.25	31.30	31.35	31.40	31.45	31.50	31.55	31.60	31.65	31.70	31.75	31.80	31.85	31.90	31.95	32.00	32.05	32.10	32.15	32.20	32.25	32.30	32.35	32.40	32.45	32.50	32.55	32.60	32.65	32.70	32.75	32.80	32.85	32.90	32.95	33.00	33.05	33.10	33.15	33.20	33.25	33.30	33.35	33.40	33.45	33.50	33.55	33.60	33.65	33.70	33.75	33.80	33.85	33.90	33.95	34.00	34.05	34.10	34.15	34.20	34.25	34.30	34.35	34.40	34.45	34.50	34.55	34.60	34.65	34.70	34.75	34.80	34.85	34.90	34.95	35.00	35.05	35.10	35.15	35.20	35.25	35.30	35.35	35.40	35.45	35.50	35.55	35.60	35.65	35.70	35.75	35.80	35.85	35.90	35.95	36.00	36.05	36.10	36.15	36.20	36.25	36.30	36.35	36.40	36.45	36.50	36.55	36.60	36.65	36.70	36.75	36.80	36.85	36.90	36.95	37.00	37.05	37.10	37.15	37.20	37.25	37.30	37.35	37.40	37.45	37.50	37.55	37.60	37.65	37.70	37.75	37.80	37.85	37.90	37.95	38.00	38.05	38.10	38.15	38.20	38.25	38.30	38.35	38.40	38.45	38.50	38.55	38.60	38.65	38.70	38.75	38.80	38.85	38.90	38.95	39.00	39.05	39.10	39.15	39.20	39.25	39.30	39.35	39.40	39.45	39.50	39.55	39.60	39.65	39.70	39.75	39.80	39.85	39.90	39.95	40.00	40.05	40.10	40.15	40.20	40.25	40.30	40.35	40.40	40.45	40.50	40.55	40.60	40.65	40.70	40.75	40.80	40.85	40.90	40.95	41.00	41.05	41.10	41.15	41.20	41.25	41.30	41.35	41.40	41.45	41.50	41.55	41.60	41.65	41.70	41.75	41.80	41.85	41.90	41.95	42.00	42.05	42.10	42.15	42.20	42.25	42.30	42.35	42.40	42.45	42.50	42.55	42.60	42.65	42.70	42.75	42.80	42.85	42.90	42.95	43.00	43.05	43.10	43.15	43.20	43.25	43.30	43.35	43.40	43.45	43.50	43.55	43.60	43.65	43.70	43.75	43.80	43.85	43.90	43.95	44.00	44.05	44.10	44.15	44.20	44.25	44.30	44.35	44.40	44.45	44.50	44.55	44.60	44.65	44.70	44.75	44.80	44.85	44.90	44.95	45.00	45.05	45.10	45.15	45.20	45.25	45.30	45.35	45.40	45.45	45.50	45.55	45.60	45.65	45.70	45.75	45.80	45.85	45.90	45.95	46.00	46.05	46.10	46.15	46.20	46.25	46.30	46.35	46.40	46.45	46.50	46.55	46.60	46.65	46.70	46.75	46.80	46.85	46.90	46.95	47.00	47.05	47.10	47.15	47.20	47.25	47.30	47.35	47.40	47.45	47.50	47.55	47.60	47.65	47.70	47.75	47.80	47.85	47.90	47.95	48.00	48.05	48.10	48.15	48.20	48.25	48.30	48.35	48.40	48.45	48.50	48.55	48.60	48.65	48.70	48.75	48.80	48.85	48.90	48.95	49.00	49.05	49.10	49.15	49.20	49.25	49.30	49.35	49.40	49.45	49.50	49.55	49.60	49.65	49.70	49.75	49.80	49.85	49.90	49.95	50.00	50.05	50.10	50.15	50.20	50.25	50.30	50.35	50.40	50.45	50.50	50.55	50.60	50.65	50.70	50.75	50.80	50.85	50.90	50.95	51.00	51.05	51.10	51.15	51.20	51.25	51.30	51.35	51.40	51.45	51.50	51.55	51.60	51.65	51.70	51.75	51.80	51.85	51.90	51.95	52.00	52.05	52.10	52.15	52.20	52.25	52.30	52.35	52.40	52.45	52.50	52.55	52.60	52.65	52.70	52.75	52.80	52.85	52.90	52.95	53.00	53.05	53.10	53.15	53.20	53.25	53.30	53.35	53.40	53.45	53.50	53.55	53.60	53.65	53.70	53.75	53.80	53.85	53.90	53.95	54.00	54.05	54.10	54.15	54.20	54.25	54.30	54.35	54.40	54.45	54.50	54.55	54.60	54.65	54.70	54.75	54.80	54.85	54.90	54.95	55.00	55.05	55.10	55.15	55.20	55.25	55.30	55.35	55.40	55.45	55.50	55.55	55.60	55.65	55.70	55.75	55.80	55.85	55.90	55.95	56.00	56.05	56.10	56.15	56.20	56.25	56.30	56.35	56.40	56.45	56.50	56.55	56.60	56.65	56.70	56.75	56.80	56.85	56.90	56.95	57.00	57.05	57.10	57.15	57.20	57.25	57.30	57.35	57.40	57.45	57.50	57.55	57.60	57.65	57.70	57.75	57.80	57.85	57.90	57.95	58.00	58.05	58.10	58.15	58.20	58.25	58.30	58.35	58.40	58.45	58.50	58.55	58.60	58.65	58.70	58.75	58.80	58.85	58.90	58.95	59.00	59.05	59.10	59.15	59.20	59.25	59.30	59.35	59.40	59.45	59.50	59.55	59.60	59.65	59.70	59.75	59.80	59.85	59.90	59.95	60.00	60.05	60.10	60.15	60.20	60.25	60.30	60.35	60.40	60.45	60.50	60.55	60.60	60.65	60.70	60.75	60.80	60.85	60.90	60.95	61.00	61.05	61.10	61.15	61.20	61.25	61.30	61.35	61.40	61.45	61.50	61.55	61.60	61.65	61.70	61.75	61.80	61.85	61.90	61.95	62.00	62.05	62.10	62.15	62.20	62.25	62.30	62.35	62.40	62.45	62.50	62.55	62.60	62.65	62.70	62.75	62.80	62.85
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------









IMAGE FILE = STPLAPP.IMG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CUM-ROW	CUM-COLUMN
1	2123454521	12343232321
2	245432321	1234323432321
3	123132432	232321
4	1234321	121212
5	121212121	12434321
6	1212	151212421
7	121212321	123543454521
8	1212132342421	1232434321
9	1212124321	12123231
10	12121232	1212321
11	1214132342	1232321
12	12121234321	1232321
13	12321	1232121
14	2343212342	12321
15	23432143432	13212321
16	23232123432	12323231

IMAGE FILE = SAMP1.IMG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CUM-ROW	CUM-COLUMN
1	123454321	12343232321
2	243421	12343232321
3	123434321	12343232321
4	12343421	12343232321
5	12343421	12343232321
6	123454343231	12343232321
7	123434321	123432323212
8	1234321	12343232321
9	12345432521	123234321
10	134542342121	12343231323212
11	12345432321	123432343231
12	123432321	123234321
13	23454323421	1234323432
14	1234321	12343234321
15	12343231	12323432321
16	123543434521	12343234323212

IMAGE FILE = SAMP2.IMG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CUM-ROW	CUM-COLUMN
1	123432124521	132321
2	254321232321	1232321
3	1343123231	1232321
4	123124321	123232
5	1234321232321	1232321
6	2343123232	1232321
7	123232123252	1232321
8	12343212321	23232
9	1234321	1212121
10	12342	12121
11	1234321	12121
12	12321	12121
13	123431	1212121
14	12342	12121
15	1243431	12121
16	123431	12121



IMAGE FILE = SAMPLE1.IMG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CON-ROW	CON-COLUMN
1	1212121	212434321
2	1212121	123434321
3	1212121	1243421
4	12121212	12343421
5	12121231	12343431
6	1212121	1234342
7	1212121	12343432
8	12121231	123434321✓
9	1212	1212121
10	1212	12121
11	1212	1212
12	12121	12121
13	121	12121
14	12121	12121
15	12121	12121
16	12121	1232121

IMAGE FILE = SAMPLE1.IMG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CON-ROW	CON-COLUMN
1	121212321	1234345432
2	121321	12342
3	121321	12434321
4	12121212321	1234345432
5	1212121321	12434321
6	12121212321	123434321
7	12121212321	1212434341
8	12121321	125343421
9	12121234345321	1232345431
10	121212343421	123234541
11	121212345321	13232343431
12	12121234321	123234321
13	1212345321	12324321
14	121212342	123234321
15	12121234321	12323434321
16	12121234321	123234321

IMAGE FILE = SAMPLE.IMG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CON-ROW	CON-COLUMN
1	12121232	1232321
2	121212321	12321
3	12131232	12321
4	121212321	1232321
5	12121231	1232321
6	1212123421	1232321
7	121212312	123232321✓
8	121212321	121232321✓
9	1231212321	12321
10	1212321	123234321
11	121212321	12321
12	121212321	1312321
13	121212321	1212321
14	131212321	1212321
15	12121232	1212321
16	121231	1212321

IMAGE FILE = SAMP6.1PG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CUN-ROW	CUN-COLUMN
------	---------	------------

1	12121234321	12323212
2	12121234321	1232321
3	1212123421	1232321
4	12121234321	1232321
5	1212123431	1232321
6	1213212342	1232321
7	12312123432	1232321
8	12123421	12323212
9	121234321	132321
10	12123432	123212
11	12123431	1232321
12	12121234321	1232321
13	1212123421	132321
14	12121234321	1232321
15	121234321	1232321
16	1312123432	1232321

IMAGE FILE = SAMP7.1PG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CUN-ROW	CUN-COLUMN
------	---------	------------

1	12321	1232121
2	12343212	1232121
3	1234321	1232121
4	234321	1232121
5	1234321	1232121
6	1234321	1232121
7	234321	1232121
8	232321	1232121
9	2432123432	12321
10	12431234321	12321
11	123432123432	23212
12	2431234321	2321
13	1232123421	12321
14	12312342	2321
15	1232123421	2321
16	1243212321	12321

IMAGE FILE = SAMP8.1PG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CUN-ROW	CUN-COLUMN
------	---------	------------

1	12353234342	2343212321
2	12353234321	13212321
3	12323432	12343212321
4	12323234321	234321231
5	23432343432	23212321
6	12343235435432	134321231
7	1235323434321	1243212321
8	2432323432	12323212321
9	123232123432	1232321
10	12323212342	1232342
11	123232123432	232321
12	2323231234321	232321
13	123232123431	232321
14	124323212342	2323431
15	1232123432	12323432
16	12323212342	2323421

IMAGE FILE = SAMPLED.TIF

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CUN-ROW	CUN-COLUMN
1	2454321	1232321
2	2421	1234323
3	251232	25252
4	123432	121
5	1212121	24341
6	1212	1212
7	12121321	12534541
8	121212342	1232421
9	121232	121
10	121232	2321
11	1212342	1232321
12	12121232	132321
13	2321	1232121
14	2512342	2321
15	252542	1321231
16	252521232	3231

IMAGE FILE = SAMPLE1.TIF

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CUN-ROW	CUN-COLUMN
1	123454321	12321
2	24321	123432321
3	1234321	13452521
4	124521	123432321
5	124521	1232321
6	2432	123432321
7	124521	1232321
8	14321	123432321
9	2545421	12323432
10	154542	1234323431
11	2421	123234321
12	254521	123234321
13	254542321	1234323432
14	154521	12343232
15	254521	1525452
16	154521	154525451

IMAGE FILE = SAMPLE2.TIF

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CUN-ROW	CUN-COLUMN
1	1543212521	152521
2	24321232	25251
3	1312321	123232
4	123132	25252
5	152125	323
6	31252	25252
7	1521252	1232321
8	131232	25252
9	1232	12121
10	12342	12121
11	123	12121
12	123	12121
13	1231	12121
14	12342	12121
15	1231	12121
16	1231	12121

IMAGE FILE = SAMPLE1.TIF  
 CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CON-ROW	CUT-COLUMN
1	1212121	243434
2	1212121	3434
3	1212121	24342
4	12121	13421
5	1212121	34341
6	1212121	14342
7	1212121	343
8	12121	13434321
9	1212	12121
10	1212	1212
11	1212	1212
12	12121	12121
13	121	212
14	12121	12121
15	1212	1212
16	12121	112

IMAGE FILE = SAMPLE1.TIF  
 CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CON-ROW	CUT-COLUMN
1	1212321	124343
2	1321	12342
3	1321	124342
4	1321	124342
5	1321	123431
6	12321	1243421
7	12321	124341
8	1321	1234542
9	1212342	123243431
10	1234342	123234541
11	121212341	12323431
12	12123432	12323431
13	12432	1232432
14	1212342	123234321
15	1212343	1232432
16	12121234321	12323421

IMAGE FILE = SAMPLE1.TIF  
 CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CON-ROW	CUT-COLUMN
1	12121232	132321
2	12121231	12321
3	121232	12321
4	121232	32321
5	121231	132321
6	12123	12321
7	12123	1321
8	121232	12321
9	121231	12321
10	12123	12321
11	1231	12321
12	121231	12321
13	12123	1212321
14	123	12321
15	12123	1212321
16	12123	12321

IMAGE FILE = SAMPO.IMG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CON-ROW	CON-COLUMN
1	121212343	1232321
2	1212342	1232321
3	1212342	1232321
4	1234	1232321
5	12123431	1232321
6	1212342	1232321
7	121212342	1232321
8	1212342	1232321
9	12123431	1321
10	1212342	12321
11	123431	1232321
12	1212342	1232321
13	1212342	132321
14	1212342	1232321
15	123431	1232321
16	12342	132321

IMAGE FILE = SAMP7.IMG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CON-ROW	CON-COLUMN
1	2321	1232121
2	234321	1232121
3	1234321	1232121
4	2321	1232121
5	2321	1232121
6	234321	1232121
7	234321	1232121
8	2321	1232121
9	2312342	321
10	24312342	321
11	24312342	2321
12	24312343	321
13	23212342	2321
14	312342	2321
15	2312342	2321
16	2431232	2321

IMAGE FILE = SAMP8.IMG

CONDENSED ROW AND COLUMN VALUES FOR 16 CHARACTERS

CHAR	CON-ROW	CON-COLUMN
1	23234342	23212321
2	2323431	13212321
3	2323432	3212321
4	1232342	2321231
5	243234342	23212321
6	13234342	132131
7	232342	2321231
8	2432342	1321231
9	232321232	132321
10	2323212342	13232
11	2321232	232321
12	232321232	23231
13	1321231	23232
14	23212342	3231
15	12321232	23232
16	232321232	23231