A Detailed Review of Feature Extraction in Image Processing Systems

2 AUTHORS:

Gaurav Kumar
Panipat Institute Of Engineering & Technol…

15 PUBLICATIONS   **12** CITATIONS

SEE PROFILE

Pradeep Kumar Bhatia
Guru Jambheshwar University of Science &…

47 PUBLICATIONS   **141** CITATIONS

SEE PROFILE

A Detailed Review of Feature Extraction in Image Processing Systems

Gaurav Kumar,
Department of Information Technology,
Panipat Institue of Engg. & Technology,
Panipat, Haryana, India
er.gkgupta@gmail.com

Pradeep Kumar Bhatia
Department of Computer Science & Engineering,
G. J. University of Science & Technology,
Hisar, Haryana, India
pkbhatia.gju@gmail.com

Abstract—**Feature plays a very important role in the area of image processing. Before getting features, various image preprocessing techniques like binarization, thresholding, resizing, normalization etc. are applied on the sampled image. After that, feature extraction techniques are applied to get features that will be useful in classifying and recognition of images. Feature extraction techniques are helpful in various image processing applications e.g. character recognition. As features define the behavior of an image, they show its place in terms of storage taken, efficiency in classification and obviously in time consumption also. Here in this paper, we are going to discuss various types of features, feature extraction techniques and explaining in what scenario, which features extraction technique, will be better. Hereby in this paper, we are going to refer features and feature extraction methods in case of character recognition application.**

Keywords— **Feature Extraction, Image Processing, Optical Character Recognition (OCR), Pattern Recognition**

I. INTRODUCTION

Feature extraction describes the relevant shape information contained in a pattern so that the task of classifying the pattern is made easy by a formal procedure. In pattern recognition and in image processing, feature extraction is a special form of dimensionality reduction. The main goal of feature extraction is to obtain the most relevant information from the original data and represent that information in a lower dimensionality space. When the input data to an algorithm is too large to be processed and it is suspected to be redundant (much data, but not much information) then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input Pattern recognition is an emerging field of research in the area of image processing. It has been used in many applications such as character recognition, document verification, reading bank deposit slips, extracting information from cheques, applications for credit cards, health insurance, loan, tax forms, data entry, postal address reading, check sorting, tax reading, script recognition etc. Character

recognition is also applicable in newly emerging areas, such as development of electronic libraries, multimedia database, and systems which require handwriting data entry. In the late 60's, these systems were still very expensive, and therefore could only be used by large companies and government agencies. Today, pattern recognition systems are less expensive. Several research works have been done to evolve newer techniques and methods that would reduce the processing time while providing higher recognition accuracy. The widely used feature extraction methods are Template matching, Deformable templates, Unitary Image transforms, Graph description, Projection Histograms, Contour profiles, Zoning, Geometric moment invariants, Zernike Moments, Spline curve approximation, Fourier descriptors, Gradient feature and Gabor features. As an example, OCR is the process of converting scanned images of machine printed or handwritten text into a computer processable format. The process of optical character recognition has following three stages: Preprocessing, Feature extraction, Classification. Hereby we are going to discuss 2nd stage i.e. Feature extraction in detail.

II. FEATURE EXTRACTION

Feature extraction is done after the preprocessing phase in character recognition system. The primary task of pattern recognition is to take an input pattern and correctly assign it as one of the possible output classes. This process can be divided into two general stages: Feature selection and Classification. Feature selection is critical to the whole process since the classifier will not be able to recognize from poorly selected features. Criteria to choose features given by Lippman are:

"Features should contain information required to distinguish between classes, be insensitive to irrelevant variability in the input, and also be limited in number, to permit, efficient computation of discriminant functions and to limit the amount of training data required"

Feature extraction is an important step in the construction of any pattern classification and aims at the extraction of the relevant information that characterizes each class. In this process relevant features are extracted from objects/ alphabets to form feature vectors. These feature vectors are then used by classifiers to recognize the input unit with target output unit. It becomes easier for the classifier to classify between different classes by looking at these features as

CPS
Conference Publishing Services

it allows fairly easy to distinguish. Feature extraction is the process to retrieve the most important data from the raw data. Feature extraction is finding the set of parameter that define the shape of a character precisely and uniquely. In feature extraction phase, each character is represented by a feature vector, which becomes its identity. The major goal of feature extraction is to extract a set of features, which maximizes the recognition rate with the least amount of elements and to generate similar feature set for variety of instance of the same symbol. The widely used feature extraction methods are Template matching, Deformable templates, Unitary Image transforms, Graph description, Projection Histograms, Contour profiles, Zoning, Geometric moment invariants, Zernike Moments, Spline curve approximation, Fourier descriptors, Gradient feature and Gabor features.

A. Importance of feature extraction

When the pre-processing and the desired level of segmentation (line, word, character or symbol) has been achieved, some feature extraction technique is applied to the segments to obtain features, which is followed by application of classification and post processing techniques. It is essential to focus on the feature extraction phase as it has an observable impact on the efficiency of the recognition system. Feature selection of a feature extraction method is the single most important factor in achieving high recognition performance. Feature extraction has been given as "extracting from the raw data information that is most suitable for classification purposes, while minimizing the within class pattern variability and enhancing the between class pattern variability". Thus, selection of a suitable feature extraction technique according to the input to be applied needs to be done with utmost care. Taking into consideration all these factors, it becomes essential to look at the various available techniques for feature extraction in a given domain, covering vast possibilities of cases.
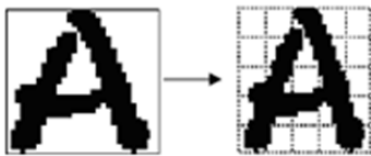

Figure 1. An image before Zoning (Feature Extraction) and after zoning

III. FEATURE SELECTION

There are two different approaches to obtain a subset of features: feature extraction and feature selection. In feature extraction the features that may have discriminating power were extracted, while in feature selection, a subset of the original set of features is selected. The main idea of features selection is to select a subset of input variables by cutout features with weakly or no predictive information while maintaining or performing classification accuracy. John et al. described feature relevance as strong and weak relevance. Strong relevance means that a feature cannot be removed from the feature set without loss of classification accuracy. Weak relevance means that a feature can sometimes contribute to classification accuracy [4].

A. Feature Selection Problem

Selecting the most meaningful features is a crucial step in the process of classification problems especially in handwriting identification because: (1) it is necessary to find all possible feature subsets that can be formed from the initial set which result in time consuming, (2) every feature is meaningful for at least some of discriminations, and (3) variations within intraclass and between inter-class is not too much high. Beyond a certain point, the inclusion of additional features leads to a worse rather than better performance.

B. Features

A good feature set contains discriminating information, which can distinguish one object from other objects. It must be as robust as possible in order to prevent generating different feature codes for the objects in the same class. The selected set of features should be a small set whose values efficiently discriminate among patterns of different classes, but are similar for patterns within the same class. Features can be classified into two categories:

1. Local features, which are usually geometric (e.g. concave/convex parts, number of endpoints, branches, Joints etc).
2. Global features, which are usually topological (connectivity, projection profiles, number of holes, etc) or statistical (invariant moments etc.)

1) Character level features: Computerized handwriting identification consist of two type of handwriting features macro and micro. Macro features are gray-scale based (entropy, threshold, no. of black pixels), contour based (external and internal contours), slope-based (horizontal, positive, vertical and negative), stroke-width, slant and height.
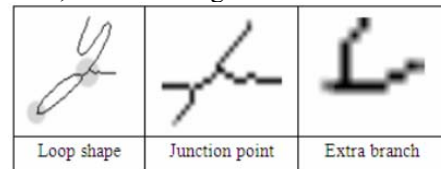

Figure 2. An example of loop shape, junction point and end

The structural features representing the coarser shape of the character capture the presence of corners, diagonal lines, and vertical and horizontal lines in the gradient image. The concavity features capture the major topological and geometrical features including direction of bays, presence of holes, and large vertical and horizontal strokes. Micro features (GSC) are found to be discriminating for the writers. Furthermore, they found 'G', 'b', 'N', 'I', 'K', 'J','W', 'D', 'h', 'f' are the 10 most discriminating characters.
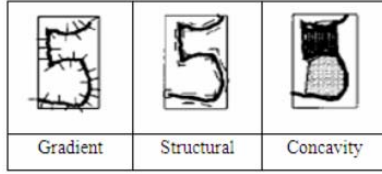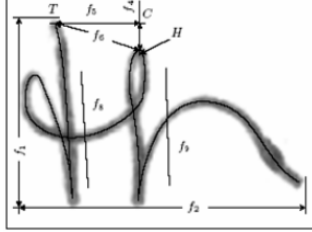
Figure 3. GSC features maps

Figure 4. Example of some geometric features extracted from grapheme 'th'

A large amount of individual information is associated with size and shape of the characters. Handwritten text has different appearances and physical structures that reflect differences of the handwriting. The features are: aspect ratio, number of end points, number of junctions, shape size and number of loops, width and height distributions, slant, shape, average curvature, and gradient features. The Hamming distance is used as classifier after these features were appropriately binarized.

IV. CLASSIFICATION OF FEATURE EXTRACTION

Feature extraction methods are classified into three major groups as:

A. Statistical Features

These features are derived from the statistical distribution of points. They provide high speed and low complexity and take care of style variations to some extent. They may also be used for reducing the dimension of the feature set. The followings are the major statistical features:

1) Zoning: The frame containing the character is divided into several overlapping or non overlapping zones and the densities of the points and some strokes in different regions are analyzed and form the features. Contour direction features measure the direction of the contours of the characters. Another example is bending point features Bending points are points at which a stroke in the image has a strong curvature.
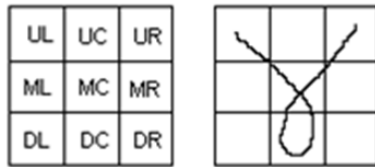

Figure 5. Zoning and x non uniform matrix representation U: Up, D: Down, M: Middle, C: Center, R: Right, L: Left

2) Characteristic Loci: For every white point in the background of the character, vertical and horizontal vectors are generated, the number of times that the line segments intersected by these vectors are used as features.

3) Crossing and Distances: Crossing counts the number of transitions from background to foreground pixels along vertical and horizontal lines through the character image. Distances calculate the distances of the first image pixel detected from the upper and lower boundaries of the image along the horizontal lines.

B. Global Transformation and Series Expansion Features

These features are invariant to global deformations like translation and rotations. A continuous signal generally contains more information that needs to be represented for the purposes of classification. One way to represent a signal is by a linear combination of a series of simpler well defined functions. The coefficients of the linear combination provide a compact encoding known as series expansion. Common transform and series expansion features are:

1) Fourier Transforms: The general procedure is to choose magnitude spectrum of the measurement vector as the features in an n dimensional Euclidean space. One of the most attractive properties of the Fourier Transform is the ability to recognize the position shifted characters when it observes the magnitude spectrum and ignores the phase. Fourier Transforms has been applied to OCR in many ways.

2) Walsh Hadamard Transform: This feature is more suitable in high speed processing since the arithmetic computation involves only addition and subtraction. The major drawback of Walsh Hadamard transform is that its performance depends heavily upon the position of the characters.

3) Rapid transform: It is same as the Hadamard Transform except for the absolute value operation which may be credited with the elimination of the position shifting problem.

4) Hough Transform: It is a technique for baseline detection in documents. It is also applied to characterize parameter curves of characters.

5) Gabor Transform: The Gabor transform is a variation of the windowed Fourier Transform. In this case the window used is not a discrete size but is defined by a Gaussian function.

6) Wavelets: Wavelet transformation is a series expansion technique that allows us to represent the signal at different levels of resolution.

7) Karhunen Loeve Expansion: It is an Eigen vector analysis which attempts to reduce the dimension of the feature set by creating new features that are linear combinations of the original features.

8) Moments: Moment normalization strives to make the process of recognizing an object in an image size translation and rotation independent.

C. Geometrical and Topological Features

These features may represent global and local properties of characters and have high tolerances to distortions and style variations. These topological features may encode some knowledge about the contour of the object or may require some knowledge as to what sort of components make up that object.

1) Strokes: These are the primitive elements which make up a character. The strokes can be as simple as lines l and arcs c which is the main strokes of Latin characters and can be as complex as curves and splines making up Arabic characters. In on line character recognition a stroke is also defined as a line segment from pen down to pen up.

Figure 6. Main strokes used for Latin characters

2) Stroke Directions and Bays: The sequence of directions of pen motion during the writing of a character is used as features.

3) Chain Codes: This feature is obtained by mapping the strokes of a character into a dimensional parameter space which is made up of codes as shown in fig. 7.

4) End points intersections of line segments and loops.

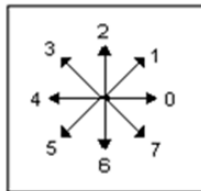*5)* Strokes relations and angular properties.

Figure 7: Chain Codes

V. APPROACHES TO DIFFERENT FEATURE EXTRACTION METHODS

Various types of feature extraction methods are shown in table I.

TABLE I. OVERVIEW OF FEATURE EXTRACTION METHODS FOR THE VARIOUS REPRESENTATION FORMS (GREY LEVEL, BINARY, VECTOR).

Grey scale subimage	Binary		Vector (skelton)
	Solid character	Outer contour	
Template matching	Template matching		Template matching
Deformable templates			Deformable templates
Unitary Transforms	Unitary Transforms		Graph description
	Projection Histogram	Contour profiles	Discrete features
Zoning	Zoning	Zoning	Zoning
Geometric moments	Geometric moments	Spline curve	
Zernike moments	Zernike moments	Fourier descriptors	Fourier descriptors

A. Diagonal based feature extraction technique

Every character image of size 90x 60 pixels is divided into 54 equal zones, each of size 10x10 pixels as shown in figure 8. The features are extracted from each zone pixels by moving along the diagonals of its respective 10 x 10 pixels. Each zone has 19 diagonal lines. Each diagonal line is summed to get a single sub-feature, thus 19 sub-features are obtained from the each zone. These 19 sub-feature values are averaged to form a single feature value and placed in the corresponding zone as shown in figure 8 (b).

This procedure is sequentially repeated for the all the zones. There could be some zones whose diagonals are empty of foreground pixels. The feature values corresponding to these zones are zero. Finally, 54 features are extracted for each character. In addition, 9 and 6 features are obtained by averaging the values placed in zones rowwise and columnwise, respectively. As result, every character is represented by 69, that is, 54 +15 features.
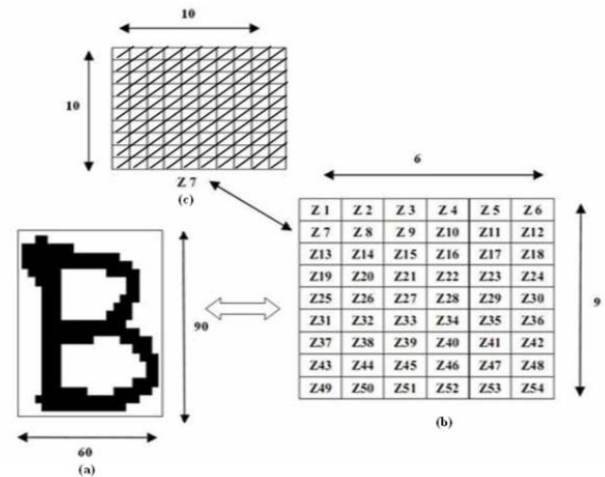

Figure 8. Procedure for extracting features from characters

Two approaches with three different ways of feature extraction are used for character recognition. The three different ways of feature extraction are horizontal direction, vertical direction and diagonal direction. Recognition rate percentage for vertical, horizontal and diagonal based feature extraction using feed forward back propagation neural network as classification phase are 92.69, 93.68, 97.80 respectively.

B. Fourier descriptor

Fourier transformation is widely used for shape analysis. The Fourier transformed coefficients form the Fourier descriptors of the shape. These descriptors represent the shape in a frequency domain. The lower frequency descriptors contain information about the general features of the shape, and the higher frequency descriptors contain information about finer details of the shape. Although the number of coefficients generated from the transform is usually large, a subset

of the coefficients is enough to capture the overall features of the shape.

Suppose that the boundary of a particular shape has K pixels numbered from 0 to K - 1. The k-th pixel along the contour has position (xk, yk). Therefore, we can describe the shape as two parametric equations:

$$x(k) = x_{k,} \qquad y(k) = y_k$$

We consider the (x, y) coordinates of the point not as Cartesian coordinates but as those in the complex plane by writing

$$s(k) = x(k) + I\, y(k)$$

We take the discrete Fourier Transform of this function to end up with frequency spectra. The discrete Fourier transform of s(k) is

$$a(u) = \frac{1}{k} \sum_{k=0}^{K-1} s(k) e^{-j2\pi uk}/_K, \qquad u = 0,1,\dots K-1$$

The complex coefficients a(u) are called the Fourier descriptors of the boundary. The inverse Fourier transform of these coefficients restores s(k). That is,

$$s(k) = \sum_{u=0}^{K-1} a(u) e^{j2\pi uk}/_K, \qquad k = 0,1,\dots K-1$$

C. Principal component analysis (PCA)

PCA is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components.

Center the data $\overline{X}$

Compute the covariance matrix C

Obtain the eign vectors and eign values of the covariance matrix U, P

Project the original data in the eigenspace $P = U^T.\overline{X}$

The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has as high a variance as possible (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to (uncorrelated with) the preceding components. Steps to compute PCA transformation of a data matrix X.

Principal components are guaranteed to be independent only if the data set is jointly normally distributed. PCA is sensitive to the relative scaling of the original variables. Depending on the field of application, it is also named the discrete Karhunen–Loeve transform (KLT), the Hotelling transform or proper orthogonal decomposition (POD).

D. Independent Component Analysis (ICA)

ICA is a statistical technique that represents a multidimensional random vector as a linear combination of nongaussian random variables (independent components) that are as independent as possible. ICA has many applications in data analysis, source separation, and feature extraction.
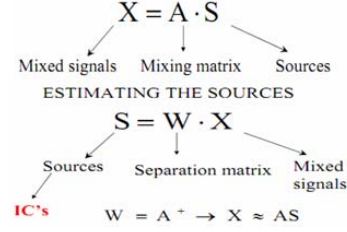
E. Gabor filter

Gabor filter possess optimal localization properties in both spatial and frequency domain. Gabor filter gives a chance for multi-resolution analysis by giving coefficient matrices. In this approach, a 2D Gabor filter gives an extracted feature. A Gabor is Gaussian modulated sinusoid in the spatial domain and as a shifted Gaussian in frequency domain. It can be represented by:

$$g_{\gamma,\eta,\varphi,\lambda} = exp\left(\frac{x'^2 + \gamma 2 y'^2}{2\sigma^2}\right) . cos\left(\frac{2\pi x'}{\lambda} + \varphi\right) \tag{1}$$

$$x' = x cos\theta - y sin\theta \tag{2}$$

$$\iint I(\varepsilon, \eta) g(x - \varepsilon, y - \eta)\, d\varepsilon d\eta \tag{3}$$

The Gabor filter can be better used by varying the parameters like λ, γ, and θ. In equation (2), x and y are image coordinates. λ is the wavelength of cosine equation, γ characterizes the shape of Gaussian, γ= 1 for circular shape, γ<1 for elliptical shape. θ represents the channel orientation and takes values in the interval (0,360). The response of Gabor filter is convolution given by equation (3).

F. Fractal theory technique

This feature can be applied to extract the feature of two-dimensional objects. It is constructed by a hybrid feature extraction combining wavelet analysis, central projection transformation and fractal theory. A multi resolution family of the wavelet is used to compute information conserving micro-features. A central projection method is used to reduce the dimensionality of the original input pattern. A wavelet transformation technique is employed to transform the derived pattern into a set of sub-patterns. Its fractal dimension can be computed and used as feature vector.

G. Shadow Features of character

For computing shadow features, the rectangular boundary enclosing the character image is divided into eight octants, for each octant shadow of character segment is computed on two perpendicular sides so a total of 16 shadow features are obtained. Shadow is basically the length of the projection on the sides. These features are computed on scaled image.

H. Chain Code Histogram of Character Contour

Given a scaled binary image, we first find the contour points of the character image. We consider a 3 × 3 window surrounded by the object points of the image. If any of the 4-connected neighbor points is a background point then the object point (P), as shown in figure 9 is considered as contour point.
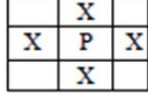
	X	
X	P	X
	X	

Figure 9. Contour point detection

The contour following procedure uses a contour representation called "chain coding" that is used for contour following proposed by Freeman, shown in figure 10a. Each pixel of the contour is assigned a different code that indicates the direction of the next pixel that belongs to the contour in some given direction. Chain code provides the points in relative position to one another, independent of the coordinate system. In this methodology of using a chain coding of connecting neighboring contour pixels, the points and the outline coding are captured. Contour following procedure may proceed in clockwise or in counter clockwise direction.
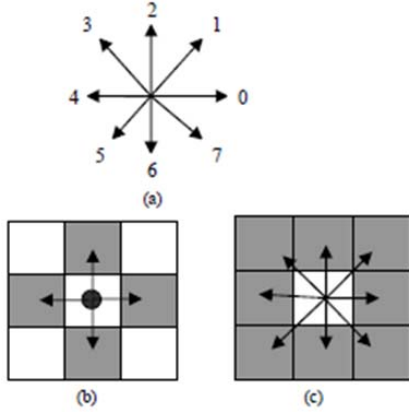

Figure 10. Chain Coding: (a) direction of connectivity, (b) 4-connectivity, (c) 8-connectivity. Generate the chain code by detecting the direction of the next line pixel

The chain code for the character contour will yield a smooth, unbroken curve as it grows along the perimeter of the character and completely encompasses the character. When there is multiple connectivity in the character, then there can be multiple chain codes to represent the contour of the character.

I. Finding Intersection/Junctions in character

An intersection, also referred to as a junction, is a location where the chain code goes in more than a single direction in an 8-connected neighborhood. Thinned and scaled character image is divided into 16 segments each of size 25 × 25 pixels wide. For each segment the number of open end points and junctions are calculated. Intersection point is defined as a pixel point which has more than two neighboring pixels in 8-connectivity while an open end has exactly one neighbor pixel. Intersection points are unique for a character in different segment. Thus the number of 32 features within the 16 constituent segments of the character image are collected, out of which first 16 feature represents the number of open ends and rest 16 features represents number of junction points within a segment. These features are observed after image thinning and scaling, as without thinning of the character image there will be multiple open end points and multiple junction points within a segment. For thinning, standard algorithm is used.
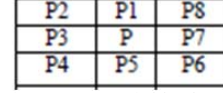
P2	P1	P8
P3	P	P7
P4	P5	P6

Figure 11. 8-neighborhoods

J. Sector approach for Feature Extraction

The recognition rate solely depends on the efficiency of features extracted from the character. Features could be topological, structural and geometrical (angles, distances). Topological or structural features work well for machine printed characters, as the shapes of these characters do not have drastic variations. However, these features alone are not suitable for recognition of hand printed characters due to some variations in writing styles. These variations could result in deformation in character shapes.
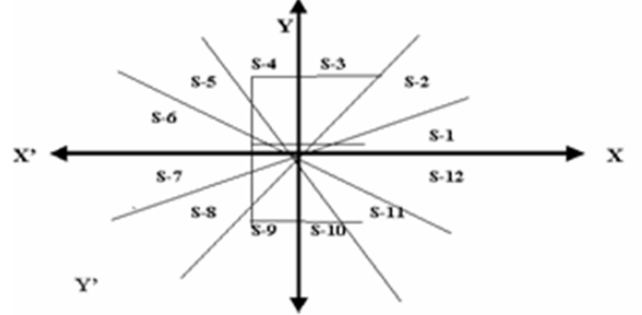

Figure 12. Formation of sectors

In the sector approach, we consider the center of the character matrix as the fixed point. This change makes the features more robust as they do not depend on the centroid. In this method, the normalized and thinned image of size (42x32) is partitioned into a fixed number of sectors from the center of the image by selecting an angle. The pictorial representation of character 'E' subdivided into 12- sectors is shown in fig. 12. The first sector is from 0 to 30 degrees; the second sector is from 30 to 60 and so on. Once the character is bifurcated into sectors, the portions lying in each sector is used for the extraction of features.

K. Extraction of distance and angle features

Let n_k be number of '1' pixels present in a sector k, with k=1,2,....,12. For each sector, the normalized vector distance, which is the sum of distances of all '1'

10

pixels in a sector divided by the number of '1' pixels present in that sector is where, (x_i , y_i) are the coordinates of a pixel in a sector and (x_m, y_n) are the co-ordinates of the center of the character image.
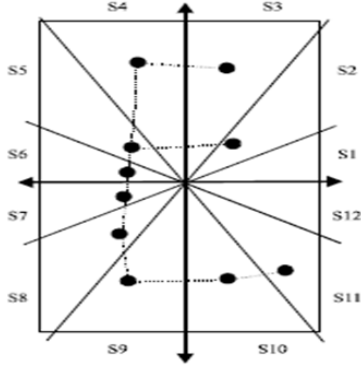

Figure 13. Reconstructed shape of 'E' using features

This normalized vector distance D_k is taken as one set of features. Next, for each sector the corresponding angles of pixels are also calculated. The normalized angle, A_k which is taken as another set of features, is calculated. Both vector distance D_k and vector angles A_k constitute 24 features from 12 sectors. These features when plotted would give an approximate pattern of the original character as shown in fig. 13.

L. Extraction of occupancy and end points features

Shape profile though depends on D_k and A_k, we augment these features proposed originally with other features such as occupancy and end points of a character for use in a recognition system. The occupancy quantifies the proportion of '1' pixels in a sector with respect to all '1' pixels in a character. We have used only four sectors for determining occupancy and accordingly. We have four values for occupancy and four values for end points. The endpoints are demarcated by tracing the character. If no '1' pixel is present, then that point is declared as an end point and sector in which it is lying is also noted as shown in fig.14. If an end point is not present in any sector, it is taken as 0 for that sector. The location of end points in a sector do not depend very much on writing styles, as they lie within a sector only. Now we have a total of 32 features consisting of 12 vector distances, 12 vector angles, 4 occupancies and 4 end points.
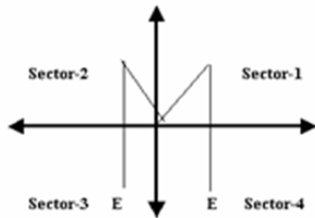

Figure 14. End-points of character 'M'

M. Transition feature

Another feature extraction approach is based on calculation and location of transition features from background to foreground pixels in the vertical and horizontal directions. Some feature extraction methods work on gray level sub-images of single characters, while others work on solid 4-connected or 8-connected symbols segmented from the binary raster image, thinned symbols or skeletons or symbol contours.

Figure 15. Gray scale subimages (=30×30 pixels) of segmented characters.

These digits were extracted from the op center portion of map in fig. 15 Note that for some of digits, part of other print objects are also present inside the character image.

Figure 16. Digit from the hydrographic map in the binary raster representation
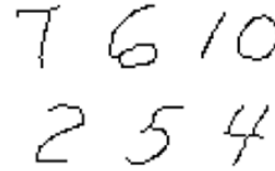

Figure 17. Skeletons of the digit in fig. 16. Junctions are displaced and a few short false branches occur.

Figure 18. Contour of two of the digit in fig.17

N. Zernike Moments

Zernike moments have been used by for character recognition of binary solid symbols. In the initial experiments, they are well suited for gray scale character subimages as well.
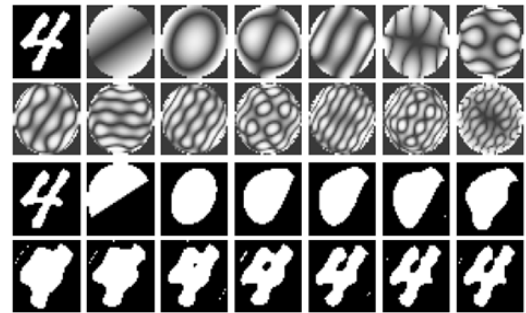

Figure 19. Images derived from Zernike moments. Row 1-2: input image of digit '4', and contributions from the Zernike moments of order 1-13.

Both rotation-variant and rotation invariant features can be extracted. Features invariant to illumination need to be developed for these features to be really useful for gray level character images. Discretization errors and other high frequency noise are removed when using Fourier descriptors, Moment invariants, or Zernike moments, since we never use very high order terms. Zoning methods are also robust against high frequency noise because of the implicit low pass filtering in the method. Of the unitary image transforms, the Karhunen loeve transform has the best information compactness in terms of mean square error. However, since the features are only linear combinations of the pixels in the input character image, we cannot expect them to be able to extract high level features thus a large training data set is needed, also since the features are tied to pixel locations; we cannot expect to get class description suitable for parametric statistical classifier. Still, a non-parametric classifier like the k-nearest neighbor classifier may perform well on the Karhunen-Loeve transform features.

VI. CONCLUSION

A detailed study about the features is discussed that may be useful in pattern recognition system. Usefulness of features, types of features in image processing system is explained. As patterns may have different orientation, styles etc., various image preprocessing techniques are applied firstly. Then based on features, some recognition technique can be applied. For extracting features, many feature extraction techniques have been developed and are discussed in this paper in detail. Using this paper, a quick overview of feature extraction techniques may be taken and it can be decided that which feature extraction technique will be better for the work to be done based on complexity, type of image (e.g. grey, color image).

REFERENCES

[1] Gaurav Y. Tawde, Jayshree Kundargi, "An Overview of features Extraction Techniques in OCR for Indian Scripts Focused on Offline Handwriting", International journal of Engineering Research and Applications, Vol. 3, Issue 1, pp. 919-926, Jan-Feb 2013.

[2] Nisha Sharma, Tushar Patnaik, Bhupender Kumar, "Recognition for Handwritten English Letters: A Review", International Journal of Engineering and Innovative Technology, Vol. 2, Issue 7, pp. 318-321, Jan. 2013.

[3] Oivind Due Trier, Anil K. Jain, Torfinn Taxt, "Feature extraction Methods for character recognition – A Survey", Journal of Pattern Recognition, Vol. 29, No. 4, pp. 641-642, 1996.

[4] Khaled Mohammed bin Abdl, Siti Zaiton Mohd, Azad Kamilah Muda, "Feature Extraction and Selection for Handwriting Identification: A Review", pp 375-381.

[5] M.Blumenstein, B. Verma, H.Basli, "A Novel Feature Extraction Technique for the Recognition of Segmented Handwritten Characters", pp.1-5.

[6] Vanita Singh, Bhupendra Kumar, Tushar Patnaik, "Feature Extraction Technique for handwritten Text in Various Script: a Survey, IJSCE, Vol. 3 , pp. 238-241, March 2013.

[7] J. Pradeep, E. Srinivasan, S.Himavathi, "Diagonal based feature Extraction for handwritten alphabets recognition system using neural network", IJCSIT, Vol. 3, No. 1, pp. 27-37, Feb. 2011.

[8] Nafiz Arica, "An Offline Character Recognition System for free style Handwriting", pp.1- 123, Sept. 2008.

[9] Anshul Gupta, Manisha Srivastava, "Offline Handwritten Character Recognition", pp. 1-27, April 2011.

[10] Vijay Laxmi Sahu, Babita Kubde, "Offline Handwritten Character Recognition Techniques using neural network: A Review", International journal of science and Research (IJSR), pp. 1-8.

[11] G. S. Lehal and Chandan Singh, "Feature Extraction and Classification for OCR of Gurumukhi Script", pp. 1-10.

[12] Vamsi K. Madasu, Brian C. Lovell, M. Hanmandlu, "Hand printed Character Recognition using Neural Networks", pp 1-10.

[13] Sandhya Arora, Meghnad Saha, Debotosh Bhattacharjee, Mita Nasipuri, "Combining Multiple Feature Extraction Techniques for Handwritten Devnagari Character Recognition", IEEE Region 10 Colloquium and the Third ICIIS, pp. 1-6, Dec. 2008

[14] Tim Klassen, "Towards Neural Network Recognition of Handwritten Arabic Letters", pp. 1-64.

[15] Cheng-Lin Liu, "Normalization-Cooperated Gradient Feature Extraction for Handwritten Character Recognition", IEEE Transaction on pattern analysis and machine intelligence, vol. 29, no. 8, Aug. 2007.

[16] Mark S. Nixon, Alberto S. Aguado, Feature Extraction and Image Processing, Newness Publisher, 2002.

[17] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, 2nd edition, Pearson Education.

[18] Gaurav Kumar, Pradeep Kumar Bhatia, Indu Banger, "Analytical Review of Preprocessing Techniques for Offline Handwritten Character Recognition", International Journal of Advances in Engineering Sciences, Vol. 3, No. 3, pp. 14-22, July 2013.

[19] Gaurav Kumar, Pradeep Kumar Bhatia, "Neural Network based Approach for Recognition of Text Images", International Journal of Computer Applications, Vol. 62, No. 14, Jan. 2013.