International Journal of Computer Theory and Engineering, Vol.3, No.1, February, 2011
1793-8201

# Feature Analysis for Handwritten Kannada Kagunita Recognition

Leena R Ragha and M Sasikumar

*Abstract*—**Handwriten Character Recognition (HCR) for Indian Languages is an important problem where there is relatively little work has been done. Particularly difficult is the problem of recognition of Kagunita – the compound characters resulting from the consonant and vowel combination. To recognize a Kagunita, we need to identify the vowel and the consonant present in the Kagunita character image. In this paper, we investigate the use of moments features on Kannada Kagunita. Kannada characters are curved in nature with some symmetry observed in the shape. This information can be best extracted as a feature if we extract moment features from the directional images. So we are finding 4 directional images using Gabor wavelets from the dynamically preprocessed original image. We analyze the Kagunita set and identify the regions with vowel information and consonant information and cut these portions from the preprocessed original image and form a set of cut images. Moments and statistical features are extracted from original images, directional images and cut images. These features are used for both vowel and consonant recognition on Multi Layer Perceptron with Back Propagation Neural Network. The recognition result for vowels are average 86% and consonants are 65% when tested on separate test data. The confusion matrices for both vowels and consonants are analyzed.**

*Index Terms*—**Gabor directional images, Handwriting Character Recognition, Moments.**

## I. Introduction

HCR is an Optical Character Recognition problem for handwritten characters. It is very valuable in terms of the variety of applications and also as an academically challenging problem. When HCR is used as a solution for inputting regional language data and also as a solution for converting paper information to soft form, HCR solutions become powerful component in addressing the digital divide. It also provides a solution to processing large volumes of data automatically. Hence extensive work is happening in this field on different scripts. But on Indian language scripts, very little work is reported and only a few research papers are available [10].

Indian scripts share a large number of structural features due to common Brahmi origin. The written form has more curves than straight or slant lines and has lots of similarities between different alphabets of the same scripts and also between the scripts of different languages. We used Kannada script for our experiments.

In an earlier paper, we had reported our results on Kannada basic character set. We explored the structural, positional, statistical and moments features extracted from the statically preprocessed images and tested on a multilayer perceptron neural network [12]. The results were average 73%. We traced one main reason for poor result as inadequate preprocessing. We formulated a new dynamic preprocessing pipeline capable of handling the variations due to pen, paper, ink color, size, ink flow variations and background noise and the results improved to average 92% [13].

Handling Kagunita (see section II) is the hardest part in Indian language HCR, since most vowels when following a consonant modifies the consonant's shape. The modifications are normally not drastic, making it hard to spot the vowel. On the other side, the modification reduces consonant recognition due to the resulting distortion in the shape. Observing that the modifications are not drastic, and that usually they are confined to some part of the letter – top, bottom, left or right, directional images appear to be a useful theme to look at for Kagunita. Using directional images and cut images with moments as features for Kagunita is the focus of this paper.

Moments are usually calculated on the preprocessed original image. Some use the moment values directly as features [1] and some find the imbalance, orientation, Hu's functions [2] etc. as features. Apart from geometric and central moments, other moments like Legendre, Zernike etc., are also explored by different researchers [7].

To find the directional information in the image, derivative operators are used by some researchers. But these are more sensitive to noise in the neighborhood [8]. Some use wavelet transformation [5][9], Gabor transformation [1] etc. Gabor features are mathematically well defined and can be fine tuned. They are less sensitive to noises and small amount of translation, rotation and scaling. Some researchers use the Gabor response of the image directly as feature [4] and some resample the image or responses and use the fewer responses as features [11][17]. Some people find the statistical features from the responses [1] and use them as features and some find the zonal statistical features of the responses [4].

In this paper, we are proposing two new concepts. The first one is a method of combining both Gabor transforms and moments and the second one is a cut images concept to extract portions of original image dominant with vowel

Manuscript received December 20, 2009.

Leena R Ragha is with the Ramrao Adik Institute of Technology (RAIT), Dr D Y Patil Vidyanagar, Sector 7 Nerul., Navi Mumbai, Maharashtra, India. (phone: 09987297843; e-mail: leena.ragha@gmail.com).

M Sasikumar is with Center for Development of Advanced Computing (CDAC), Rain Tree Marg, Sector 7, C B D Belapur, Navi Mumbai, Maharashtra, India. (phone: 09821165583 e-mail: sasi@cdacmumbai.in).

information and consonant information to recognize them separately. A neural network – multi layer perceptron (MLP) with back-propagation (BP) is used for classification. The confusion matrix is formed for analysis.

The paper is organized as follows. Section II briefly discusses Kannada script including Kagunita structure. Section III is devoted to explain briefly the technologies used for feature extraction. Section IV describes the feature extraction. Section V reports the experimental results and analysis. Section VI gives the conclusion of our study.

## II. Kannada Language Script

The current Kannada basic character set has 15 vowels and 34 consonants in the modified list. The manual observation of this set showed that many characters have similar shapes with slight modifications as shown in fig. 1 with typed characters in similar shape groups and with a sample handwritten page of basic character set. Similar shapes are marked with same colored circle and square. It is manually observed that there are about 13 groups of similar character shapes.

| | |
|---|---|
| ಅ ಆ | ಇ ಈ |
| ಐ ಣ ಟ ಚ | ಋ ಬ ಭ ಒ |
| ದ ಡ ಪ ಹ ವ ಏ | ಥ ಧ ಢ ಫ ಷ ಘ ಛ |
| ಠ ರ ಗ | ನ ಸ |
| ಳ ಶ | ಯ ಝ ಮ |
| ಕ ತ | ಙ ಜ ಓ |
| ಅಂ ಅಃ | |

Figure 1. Basic Kannada Character set with similar shape characters grouped

The Kannada Kagunita is a set of compound characters formed by combining one of the 15 vowels as in figure 2 (a) with each of the 34 basic consonants. That is when a dependent consonant combines with an independent vowel, an Akshara is formed. Thus the Kannada Kagunita has 15x34 = 510 different unique shaped Aksharas. If a character appears in its original form, it is assumed to go with the first vowel ಅ. All other vowels modify the consonant appearance in the same way. The explicit appearance of a vowel in a syllable overrides the inherent vowel ಅ of a consonant. The other vowel matra shapes and positions are as in fig. 2(b). The dotted circle indicates the position of the consonant. Sample examples of combining consonant with vowel forming a Kagunita are shown in fig. 2 (c) and (d). The first character in these examples is the basic character in its original form and the remaining 14 characters show the shape changes due to added matra.

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (a) | ಅ | ಆ | ಇ | ಈ | ಉ | ಊ | ಋ | ಎ | ಏ | ಐ | ಒ | ಓ | ಔ | ಅಂ | ಅಃ |
| (b) | | ◌ಾ | ◌ಿ | ◌ೀ | ◌ು | ◌ೂ | ◌ೃ | ◌ೆ | ◌ೇ | ◌ೈ | ◌ೊ | ◌ೋ | ◌ೌ | ◌ಂ | ◌ಃ |
| (c) | ಕ | ಕಾ | ಕಿ | ಕೀ | ಕು | ಕೂ | ಕೃ | ಕೆ | ಕೇ | ಕೈ | ಕೊ | ಕೋ | ಕೌ | ಕಂ | ಕಃ |
| (d) | ಯ | ಯಾ | ಯಿ | ಯೀ | ಯು | ಯೂ | ಯೃ | ಯೆ | ಯೇ | ಯೈ | ಯೊ | ಯೋ | ಯೌ | ಯಂ | ಯಃ |

Figure 2. (a) 15 vowels, (b) matras (c) ,(d) sample examples of Kagunita

Thus recognizing Kagunita means recognizing 510 different aksharas, a very large number of classes and so is a very complex task. Since Kagunita is a combination of vowels and consonants, if we identify these two components in a given form, then we can recognize a Kagunita. Hence the problem is now recognizing a vowel from 15 vowels and a consonant from 34 consonants extracted from a Kagunita. This is the approach we take in this paper.

The recognition of vowel in the presence of the 34 variations of consonant and the recognition of a consonant under the presence of 15 vowel variations is a difficult task. As the size of the vowel matra varies, this influences the size of the consonant in the normalized image. A vowel matra has similarities with other vowel matra and a consonant has similarities with other consonants. Hence recognition of a vowel or a consonant under the presence of huge similarities and dissimilarities makes the problem complex. This raises issues of looking at two types of feature extraction – one suitable for identifying the vowel element and another set for the consonant element.

## III. Technology for Feature Extraction

Kannada characters are curve shaped with some regions highly dense than others. Some shapes are wider and some are longer than others with some kind of symmetry observed in most cases. The complexity in forming a Kagunita, makes its recognition complex, thus one category of features cannot help in its recognition. In this paper we are exploring the use of moments as features. We are proposing a new method of finding moments as features from the four directional images obtained using Gabor wavelets. Statistical features are also considered to strengthen the feature set.

### A. *Gabor Transformation*

A Gabor filter is a linear filter whose impulse response is defined by a harmonic function multiplied by a Gaussian function. Because of the multiplication-convolution property, the Fourier transform of a Gabor filter's impulse response is the convolution of the Fourier transform of the harmonic function and the Fourier transform of the Gaussian function [16].

A 2D Gabor filter can be described by the impulse response function (IRF) (1) at a sampling point (x,y) with wavelength λ, oscillation frequency $f_0 = 1/\lambda$ and oscillation orientation θ.

$$h(x,y,\lambda,\theta) = g(x',y')e^{j2\pi(u_0x+v_0y)} \tag{1}$$

with $$g(x,y) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-1}{2}((x/\sigma)^2+(y/\sigma)^2)} \tag{2}$$

where $x' = x\cos\theta + y\sin\theta$ and $y' = -x\sin\theta + y\cos\theta$ are the rotated coordinates and $\sigma = \sigma_x = \sigma_y$ is the standard deviation along the x-direction and y-direction which decides the spread (scale) of the 2D Gaussian envelope. The spatial frequency is $u_0 = f_0\cos\theta$ and $v_0 = f_0\sin\theta$.

Some authors say that $\sigma_x$ and $\sigma_y$ can be different and are the functions of λ [3]. The IRF of Gabor filter is a complex function and comprises an even-symmetric component (real part) and an odd-symmetric component (imaginary part). The real part of the Gabor filter is as shown in the fig. 3.

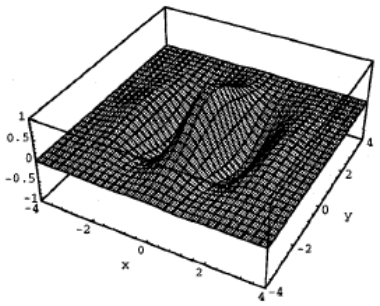



Figure 3. Real part of Gabor filter for $\lambda=5\sqrt{2}$, $\sigma=0.5*\lambda$, $\theta=45^0$

In a wavelet framework, the multiple scale parameters of Gabor filters are inter-related. The frequencies are related logarithmically, the Gaussian envelope has a constant aspect ratio and its scale is proportional to the wavelength. The scaling factor is usually selected from $v = \sqrt{2}$ (0.5 octave), 2 (1 octave) and $2\sqrt{2}$ (1.5 octave).

Accordingly, the wavelength $\lambda \propto v^j$ ($j$=0, 1, … index of the scale ), and the Gaussian function has $\sigma_x = \frac{\sigma_0}{f_0}$ with aspect ratio $\alpha = \frac{\sigma_y}{\sigma_x}$ a constant.

That is the Gabor filter has only 2 free parameters $f_0$ and θ for generating the Gabor wavelets. Gabor filters are directly related to Gabor wavelets, since they can be designed for number of dilations and rotations. However, in general, expansion is not applied for Gabor wavelets, since this requires computation of biorthogonal wavelets, which may be very time-consuming. Therefore, usually, a filter bank consisting of Gabor filters with various scales and rotations is created. The IRFs of both real and imaginary parts of Gabor filters of 2 scales and four orientations are shown in fig. 4. The first four in the first row are real parts and the next four are the imaginary parts of the Gabor filter of smaller scale with $\lambda = 1.5\sqrt{2}$. The second row is the real part and the third row is the imaginary part with bigger scale of $\lambda = 5\sqrt{2}$ [6].

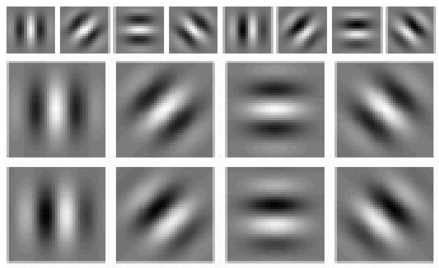


Figure 4. Real and imaginary parts of Gabor filters of 2 scales and four orientations.

An image is a 2 dimensional signal with light intensity as the magnitude of each signal; the intensity variations decide the frequency components in the image. The 2D IRF magnitude can be considered as an image. Since only one part magnitude is sufficient to find the directional response we use only real part in the computations. The DC response of the Gabor filter is dependent on the average intensity of the pixels, which depends on the illuminating conditions and thickness of the contour which are variable. A number of experiments are carried out to find the suitable λ value to fit variations of the boundary thickness. More the value of λ more the spread of the filter in the neighborhood as shown in fig. 4 and hence will merge the close by contour responses. So it is found that the thinned binary image is suitable over the original gray images of varying stroke length for the experiment.

As the Kannada character shapes are more curved than straight, we are interested in finding the density and the position of these curves. Since the shape, direction and the amount of curvature vary from one handwritten character to another, choosing more directions will give the similar IRFs as shown in fig. 5. By comparing 0 with π/8 and π/4 with 3π/8 we find that such close orientations will not help us to cover the human writing variations. By choosing proper λ, the spread of the filter can be manipulated to cover this close by orientations and hence making the filter more robust to small variations in orientation, translation and scaling [16]. Hence we are interested in only 4 major directions.

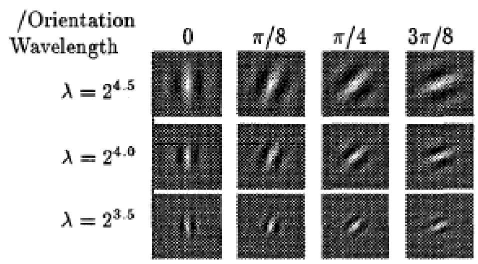



Figure 5. Real components of the Gabor filters with different λ and θ.

For a character image F(x,y), its Gabor transform result is obtained by applying Gabor filter h(x,y,λ,θ) with x Є (0..M) and y Є (0..N) to the edge image, where M and N are the dimensions of the Gabor filter. The response output can be defined through the convolution sum

$$I(x,y,\theta,\lambda) = \sum_{x1=x-M/2}^{x+M/2} \sum_{y1=y-N/2}^{y+N/2} F(x1,y1)\cdot h(x,y,\lambda,\theta) \tag{3}$$

Based on the experiments, finally L = 1 wavelength with λ

$= 5\sqrt{2}$, $\sigma = 0.5^* \lambda$ and K = 4 orientations with $\theta = \left\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\right\}$ are chosen. This constitutes a set of K×L = 4 Gabor wavelets for the experiment.

By convolving these wavelets with the preprocessed character image F(x,y), a set of convolution outputs are calculated. These outputs exhibit perfect local space characteristics, frequency characteristics and orientation selectivity of the image. The directional feature information is again converted into a binary image to extract moments features as shown in the fig. 6. The four images corresponding to one original image have only selected direction edges with the thickness of the edge depending on the scale of the Gabor filter.

$0^0$ $45^0$ $90^0$ $135^0$ Original image

Figure 6. Results of Gabor Transform on 3 images of different aspect ratio

### B. Moments

Moments are robust to high frequency noise because high order terms are not used in the feature formation, but, they require large number of computations. Carefully selected moment functions can ensure that the extracted features are invariant under translation, scale change and rotation. More importantly, moments can represent each character uniquely regardless of how close the characters are in terms of local features[2][6]. This unique nature makes moments appropriate for handwriting character recognition.

1) Geometric Moments

For a digital image f(x,y) of size MxN, the geometric moment of order (p,q) is given by

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q \; f(x,y) \;\; with \;\; p,q = 0,1,\cdots\infty \tag{4}$$

All $m_{pq}$ with p+q ≤ n a positive integer, are the geometric moments of order p+q. For a binary image, the zero order moment $m_{00}$ represents the total area of an image. The two first order moments, $m_{01}$ and $m_{10}$ represent center of mass of an image. In terms of moment values, the coordinates of center of mass also called centroid is

$$\bar{x} = \frac{m_{10}}{m_{00}} \; and \; \bar{y} = \frac{m_{01}}{m_{00}} \tag{5}$$

The second order moments $m_{11}$, $m_{20}$, $m_{02}$ give moment of inertia and they characterize the size and orientation of the image. For character recognition, normalization with respect to orientation is usually not desirable since it would result in confusion among certain classes.

2) Central Moments

To make features translation invariant, the M x N image plane is to be mapped onto a square defined by x Є [-1,+1] and y Є [-1,+1] as shown in the fig. 7.

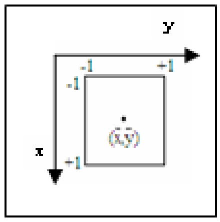



Figure 7. Normalized image coordinates

Invariance with respect to position of the object in the image can be achieved by calculating the central moments of the mapped digital image.

$$\mu_{pq} = \sum_{x=-1}^{+1} \sum_{y=-1}^{+1} (x-\bar{x})^p (y-\bar{y})^q f(x,y) \;\; with \;\; p,q = 0,1,\cdots\infty \tag{6}$$

Invariance to scale is achieved by calculating the new set of moments $\eta_{pq}$.

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{pq}^{\nu}} \; with \;\; \nu = \frac{p+q+2}{2} \tag{7}$$

Using the 0th, 1st, 2nd and 3rd order moments the following features are obtained.

4 Non linear functions using geometric moments [6]

$$f1 = m_{20} + m_{02} + m_{00}, \;\; f2 = \sqrt{(m_{20} - m_{02})^2} + m_{11}, \\ f3 = \sqrt{(m_{10} - m_{01})^2} \;\; and \;\; f4 = m_{30} + m_{03} \tag{8}$$

3 Non linear function using central moments

$$f1 = \mu_{20} + \mu_{02} + \mu_{00}, \;\; f2 = \sqrt{(\mu_{20} - \mu_{02})^2} + \mu_{11} \\ and \;\; f4 = \mu_{30} + \mu_{03} \tag{9}$$

7 non linear Hu's functions invariant to scale rotation and translation using central moments [2] are

$$\begin{aligned}
&\Phi1 = \eta_{20}+\eta_{02} \\
&\Phi2 = (\eta_{20}+\eta_{02})^2+4\eta_{11}^2 \\
&\Phi3 = (\eta_{30}-\eta_{12})^2+(3\eta_{21}+\eta_{03})^2 \\
&\Phi4 = (\eta_{30}+\eta_{12})^2+(\eta_{21}+\eta_{03})^2 \\
&\Phi5 = (\eta_{30}-3\eta_{12})(\eta_{30}+\eta_{12})[(\eta_{30}+3\eta_{p12})^2-3(\eta_{21}+\eta_{03})^2]+ \\
&\quad (3\eta_{21}-\eta_{03})(\eta_{21}+\eta_{03})[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2] \\
&\Phi6 = (\eta_{20}-\eta_{02})[(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2]+ \\
&\quad (\eta_{11}(\eta_{30}+\eta_{12})(\eta_{21}+\eta_{03}) \\
&\Phi7 = (3\eta_{12}-\eta_{30})(\eta_{30}+\eta_{12})[(\eta_{30}+\eta_{12})^2-3(\eta_{21}+\eta_{03})^2]+ \\
&\quad (3\eta_{21}-\eta_{03})(\eta_{21}+\eta_{03})[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2]
\end{aligned} \tag{10}$$

The numerical values of Φ1 to Φ7 are very small. Hence the logarithm of the absolute values of these functions is used as features. All the moment features need to be normalized before they are used as input features to Neural Network. The inter class mean and the Standard Deviation (SD) of the features are used for the normalization as given in (11)

$$Nf(x) = \frac{f(x) - mean(x)}{SD(x)} \tag{11}$$

### C. Statistical Features

Some characters have more density on the top, some in the bottom and some in the middle of the character when viewed in the horizontal direction. Similar observation results when viewed in the vertical direction. Also some have more information, some less and some close to the average information of the character with thin, wide or squared shapes. These features play an important role to find the density of the information and its distribution in the image and so we considered statistical features also to strengthen the feature set.

1) Information content of the image

A global feature called information content of the image is chosen as one of the feature since some characters have more number of curvatures making image to have more information in the same space. This feature is calculated as the ratio of the total number of black pixels to the total number of pixels in the image.

$$IC = T_{black} / T_{allpixels} \tag{12}$$

2) Aspect Ratio

Some of the characters can be easily distinguished because of their typical form, like their thinness, broad-ness, square shaped-ness or tallness etc,. The minimum and the maximum positional values of x and y are used in the computation.

$$AR = \frac{(y_{\max} - y_{\min})}{\sqrt{\left((y_{\max} - y_{\min})^2 + (x_{\max} - x_{\min})^2\right)}} \tag{13}$$

3) Zonal probability distribution

Histograms count the number of pixels in each row or column of a character. Histograms help to distinguish the characters with respect to the variation of the density of information. On cut images and on Gabor directional images, probability distribution of information is calculated using equation (14).

$$P(z) = T_{zone} / T_{black} \tag{14}$$

The whole image is divided into fixed size 5 horizontal zones 5x1 and 5 vertical zones 1x5 with a total of 10 zones. Each cut image (see section IV B) is divided into fixed size 3 horizontal 3x1 and 3 vertical 1x3 zones for zonal feature extraction resulting in 6 zones. From each zone probability distribution of information is computed.

For whole original image and four directional images, we find information content of the image, aspect ratio and the zonal probability distribution. For each of the cut images we find only zonal probability distribution.

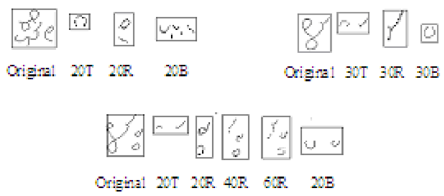
### D. *Cut images for Kagunita*

Kannada script does not have any vowel that modifies the left side of the consonant. The vowel sign is on the top, right or bottom of the character. The matra are grouped based on the top variations, side variations and bottom variations as shown in the fig. 8.

There are very small variations of matra shapes with same size on the top. In bottom there are only two variations of same size. To make this information prominent for its recognition, this portion can be cut around 20% on the top and bottom respectively giving 2 different images. The right vowel matra varies in size. We can observe four different sizes of vowel matra.

To recognize the right side vowel, we therefore considered 4 different size images - one original image and 3 cut images so that at least one category can best represent the vowel matra. The first image category is the one with no cuts (original whole image), second category with around 20% cuts, third with around 40% cut and the fourth category with around 60% cut. Hence there will be around 6 images to help in the recognition of vowel in the Kagunita as shown in the fig. 9(a).

Similarly, to recognize the consonant component in the Kagunita, as the images are size normalized, the vowel matra variation varies the size of the consonant present to the left of the vowel matra. Hence if no vowel matra to the right then consonant occupies the whole image. With 20% vowel matra the consonant occupies 80% of the total space and so on. Therefore to recognize the consonant we considered again 4 images with one original and 3 cut images as shown in the fig. 9(b).




(a) cut images for vowel recognition


(b) cut images for Consonant recognition

Figure 8. % cuts on the original images for vowel and consonant information

## IV. Feature Extraction

To recognize the Kagunita, both vowel and consonant are to be recognized. The problem becomes complicated since separating of vowel and consonant information from a given handwritten Kagunita image is very difficult due to high writing variations and hence there is a need for a very robust set of features. We are exploring different methods of extracting and combining features to address these problems.

### A. *Moment features from Original and Directional images*

Kannada character's curvature shape can be best extracted as a feature if we extract moment features from the directional images. So we are finding 4 directional binary images using Gabor wavelets from the dynamically preprocessed original images. We then extract moments features from them. The directional image centralness, divergence, imbalance should describe the relevant characteristics of the original image much more strongly than if moments are taken on the original image.

Based on this intuition, we are computing the moments features – 4 geometric moments, 10 central moments with 3 non linear and 7 Hu's moments to find centralness, divergence and imbalance, from both original images and the directional images for the analysis. The geometric moments, central moments and the combined moments from the original images are abbreviated as O-GM, O-CM and O-GMCM respectively. Similarly the geometric moments, central moments and the combined moments from the directional images are abbreviated as G-GM, G-CM G-GMCM respectively. To strengthen the feature set, we considered statistical features. The zonal probability distribution of information, aspect ratio and information content from four directional images froms the feature set G-S.

### *B. Features from cut images for vowel and consonant recognition*

To see the effect of loosing some portion of the character shape, we performed experiments on basic character set with 200 samples each of 49 characters (15 vowels and 34 consonants). The preprocessed images are cut by 20% from all the directions and the structural, positional, statistical and moments features (SPM) are extracted [13]. The results of training and testing of these features on MLP BP neural network is as in table I. It is observed that the results are down by 4% only. Based on these experiments we proposed cut images concept for Kagunita recognition.

TABLE I. RECOGNITION RESULTS OF BASIC CHARACTER SET

| Sl No | No of features | Basic character set | | | |
|---|---|---|---|---|---|
| | | Feature description | Recognition results in % on test data | | |
| | | SPM features from | Mean | Min | Max |
| 1 | 94 | original images | 93 | 83 | 100 |
| 2 | 94 | 20% cut original images | 89 | 75 | 100 |

To recognize the vowel component of the Kagunita, the original image is cut by some percentage from top, right and bottom directions. As there are 4 different vowel sizes to the right of the consonant, we consider one cut or three cuts to the right to check the effect of 3 cuts over single cut. The % cut image sizes considered for the vowel recognition experiments are as listed in table II. Since the cuts are done based on manual observations, a number of trials are done to find the optimal cut positions on the original image for both vowel and consonant recognition. Here 20T20R20B means feature set is formed from the cut images- 20% top, 20% right and 20% bottom.

TABLE II. SIZE AND COUNT OF CUT IMAGES FOR VOWEL RECOGNITION

| Sl No | Vowel Feature description | Top | | Right | | Bottom | |
|---|---|---|---|---|---|---|---|
| | | No of Images | % cut | No of Images | % cut | No of Images | % cut |
| 1 | 20T20R20B | 1 | 20 | 1 | 20 | 1 | 20 |
| 2 | 30T30R30B | 1 | 30 | 1 | 30 | 1 | 30 |
| 3 | 30T50R30B | 1 | 30 | 1 | 50 | 1 | 30 |
| 5 | 20T204060R20B | 1 | 20 | 3 | 20,40,60 | 1 | 20 |
| 6 | 20T305070R20B | 1 | 20 | 3 | 30,50,70 | 1 | 20 |

For Consonant recognition, similar experiments are done to find the most appropriate 3 cuts from the left on the original image to get 3 separate images. Since the bottom shape of the character is very significant to distinctly identify the compound characters, one more image is obtained from this region for the experiments and trials are done with and without it to check its contribution. The % cut image sizes considered for the consonant recognition experiments are as listed in table III.

From all the cut images, the geometric moments, central moments and the statistical features are computed. The feature set also includes the zonal probability distribution of information.

TABLE III. SIZE AND COUNT OF CUT IMAGES FOR CONSONANT RECOGNITION

| Consonant feature description | Bottom | | Left | |
|---|---|---|---|---|
| | No of Images | % cut | No of Images | % cut |
| 507090 | - | - | 3 | 50,70,90 |
| 406080 | - | - | 3 | 40,60,80 |
| 305070 | - | - | 3 | 30,50,70 |
| 305070L30B | 1 | 30 | 3 | 30,50,70 |
| 406080L30B | 1 | 30 | 3 | 40,60,80 |

## V. RESULTS AND ANALYSIS

We compiled a corpus of 15 samples each of 510 Kannada Kagunita characters with no restriction on the pen, paper, and ink color, ink flow, size etc. Hence the corpus has 7650 images. We proposed a dynamic preprocessing pipeline that is capable of handling all these variations with categorization of images [13]. The dynamic preprocessing decides the preprocessing stages to be applied over the input image based on 2 factors. The first factor is the foreground intensity distribution (FID) and background intensity distribution (BID) and the second factor is the ink spread and the original size of the image with background. It converts the original color / gray image into a thinned, size normalized, boundary touching, noiseless character image with aspect ratio of the original shape preserved. These fixed size images are used for feature extraction.

A number of experiments are performed to check the individual feature set performance. We worked on moments as main features from original images and Gabor directional images of the basic character set and found that they offer complimenting discrimination powers [14]. We tried these techniques for Kagunita recognition and as expected the results showed complimenting discrimination powers [15].

For Kagunita recognition, the experiments are done by considering 70% of the total sample as train data and 30% as test data. For consonants, with 7650 images 225 sets of samples is formed with 34 consonants features in each set. For the experiment, 57 sets of samples are formed as test data from 225 sets of samples. Similarly for vowels, with 7650 images, 510 sets are formed. From these, 128 sets of samples are separated out as test samples.

*A. Performance of moments and statistical features on whole Kagunita images*

For the experiments we considered features from original images and directional images. From dynamically preprocessed original images we found Central moments and Geometric moments and to test the efficiency of the feature set we considered Central moments from original images as one feature set O-CM and both central and geometric moments O-GMCM as another feature set to find their contributions, and both feature sets are normalized for vowel and consonant recognition. The recognition capabilities are tested and results of vowel recognition are as in table IV and of consonant are as in table V. Geometric moments improved the results from 40% to 47% for vowels and 24% to 29% for consonants.

TABLE IV. VOWEL RECOGNITION RESULTS OF MOMENTS FEATURES FROM WHOLE IMAGES

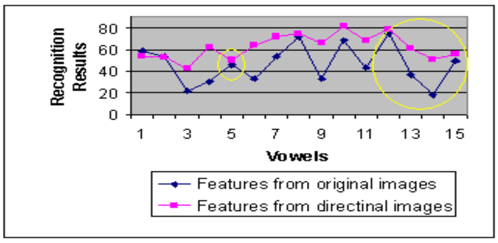
| Sl no | No of features | Feature description | Vowel Recognition results in % No of epocs – 1000. | | |
|---|---|---|---|---|---|
| | | | Mean | Min | Max |
| 1 | 10 | O-CM | **40** | 15 | 70 |
| 2 | 14 | O-GMCM | 47 | 18 | 74 |
| 3 | 40 | G-CM | **63** | 42 | 81 |
| 4 | 56 | G-GMCM | 64 | 40 | 82 |

Similarly we computed central moments G-CM and both geometric and central moments together G-GMCM as features on four directional images and normalized for vowels and consonants recognition. The recognition capabilities are tested and results of vowels are as in table IV and of consonants are as in table V. G-GMCM contributed only 1% improvement from 63% to 64% over G-CM for both vowels and consonants.

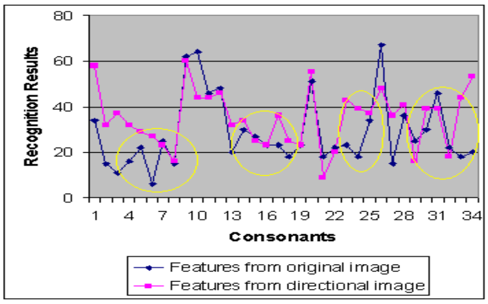
TABLE V. CONSONANT RECOGNITION RESULTS OF MOMENTS FEATURES FROM WHOLE IMAGES

| Sl no | No of featu-Res | Feature description | Consonant Recognition results in % No of epocs – 1000. | | |
|---|---|---|---|---|---|
| | | | Mean | Min | Max |
| 1 | 10 | O-CM | **24** | 8 | 51 |
| 2 | 14 | O-GMCM | 29 | 4 | 62 |
| 3 | 40 | G-CM | **35** | 9 | 60 |
| 4 | 56 | G-GMCM | 36 | 9 | 62 |

The plot of geometric and central moments from original images and central moments from directional images for 15 vowels and 34 consonants are as shown in the fig. 10 (a) and (b) respectively.




(a) Recognition results for vowels




(b) Recognition results for consonants,

Figure 9. Individual class Moments results from series 1- original image, series-2 four directional Gabor images

As expected they show a nonlinear relation as shown with yellow circles. This suggests they can be used together as features in further experiments.

The combined moments vowel recognition results along with statistical features is as shown in table VI. When moment features G-CM and O-GMCM are together used as features the vowel average recognition results improved from 63% to 71%. The statistical features G-S showed a performance of 21% individually but when combined with G-CM the results improved from 63% to 69%. When all these features here used together, the average results are 77%.

TABLE VI. VOWEL RECOGNITION RESULTS OF COMBINED MOMENTS FEATURES FROM WHOLE IMAGES

| Sl no | No of features | Feature description | Vowel Recognition results in % No of epocs – 1000. | | |
|---|---|---|---|---|---|
| | | | Mean | Min | Max |
| 1 | 54 | G-CM+O-GMCM | 71 | 57 | 83 |
| 2 | 48 | G-S | 21 | 9 | 43 |
| 3 | 88 | G-CM+G-S | 69 | 50 | 84 |
| 4 | 102 | G-CM+O-GMCM+G-S | 77 | 56 | 92 |

Table VII lists the combined moments and statistical features recognition results for consonants. When moment features G-CM and O-GMCM are together used as features the recognition result is average 44% with an improvement of 9%. The statistical features G-S did not perform well but when combined with G-CM the results improved from 35% to 43% and along with O-GMCM the result is 49%.

TABLE VII. CONSONANT RECOGNITION RESULTS OF COMBINED MOMENTS FEATURES FROM WHOLE IMAGES

| Sl n o | No of features | Feature description | Consonant Recognition results in % No of epocs – 1000. | | |
|---|---|---|---|---|---|
| | | | Mean | Min | Max |
| 1 | 54 | G-CM+O-GMCM | 44 | 22 | 65 |
| 2 | 48 | G-S | 10 | 01 | 29 |
| 3 | 88 | G-CM+G-S | 43 | 20 | 67 |
| 4 | 102 | G-CM+O-GMCM+G-S | 49 | 23 | 78 |

*B. Performance of moments and statistic features on cut images:*

A number of experiments are performed to extract the vowel information from the original image using different cut images as shown in table I. The results are as in table VIII. A

20T305070R20B cut with one image of 20% cut on the top, 3 images with 30%, 50% and 70% cut from the right and one image with 20% cut sizes on the bottom of the original image performed well for vowel recognition compared to other cuts with an average of 69%.

TABLE VIII. RESULTS OF CUT EXPERIMENTS FOR VOWEL RECOGNITION

| Sl No | No of features | Vowels | | | |
|---|---|---|---|---|---|
| | | Feature description | Recodnition results in % on test data | | |
| | | | Mean | Min | Max |
| 1 | 48 | 30T50R30B | 52 | 25 | 75 |
| 2 | 48 | 20T20R20B | 57 | 36 | 84 |
| 3 | 48 | 30T30R30B | 55 | 26 | 78 |
| 4 | 80 | 30T204060R30B | 66 | 47 | 90 |
| 5 | 80 | 20T204060R20B | 69 | 48 | 83 |
| 6 | 80 | **20T305070R20B** | **69** | **50** | **87** |

Similar experiments are done for consonant feature extraction. The different percent cuts from the left are tried using different cut images as in table II. The results are as in table IX. Experiments are done with 3 different percent cuts. As the consonant shape in the bottom portion carries distinguishing information of similar shapes, some experiments are performed by considering 3 left cut images and one bottom cut image together but, we found that the system performance reduced by 10%. A 406080 cut with 3 images of 40%, 60% and 80% cut sizes from the left performed well compared to others with an average of46%.

TABLE IX. RESULTS OF CUT EXPERIMENTS FOR CONSONANT RECOGNITION

| Sl No | No of features | Consonants | | | |
|---|---|---|---|---|---|
| | | Feature description | Recognition results in % on test data | | |
| | | | Mean | Min | Max |
| 1 | 48 | 507090 | 42 | 18 | 78 |
| **2** | **48** | **406080** | **46** | **15** | **88** |
| 3 | 48 | 305070 | 34 | 14 | 64 |
| 4 | 64 | 305070L30B | 37 | 16 | 76 |
| 5 | 64 | 406080L30B | 35 | 13 | 72 |
| 6 | 64 | 507090L30B | 36 | 11 | 74 |

The moments and statistical features extracted from cut images 20T305070R20B and from four Gabor directional images G-CM+G-S performs well with an average recognition rate of 85% for vowel recognition as in table X. For O-GMCM the result is 86% with an improvement of 1% for vowel recognition.

TABLE X. VOWEL RECOGNITION RESULTS WITH COMBINED FEATURE SET

| Sl no | No of features | Feature description | Vowel Recognition results in % No of epocs – 1000. | | |
|---|---|---|---|---|---|
| | | | Mean | Min | Max |
| 1 | 168 | G-CM+G-S+ **20T305070R20B** | 85 | 75 | 94 |
| 2 | 182 | G-CM+G-S+O-GMCM **+20T305070R20B** | 84 | 72 | 93 |

Similarly for consonants the recognition rate is average 63% as in table XI. With O-GMCM there is an improvement of 2% for consonant recognition. The maximum consonant recognition rate is 65% which is not sufficient to build a practically feasible HCR solution. To uniquely recognize the 34 different consonants under huge vowel shape variations and close consonant similarities, the feature set need to be strengthened. For this we computed the confusion matrix to analyze the reason for confusion.

TABLE XI. CONSONANT RECOGNITION RESULTS WITH COMBINED FEATURE SET

| Sl no | Noof features | Feature description | Consonant Recognition results in % No of epocs – 1000. | | |
|---|---|---|---|---|---|
| | | | Mean | Min | Max |
| 1 | 136 | G-CM+G-S+406080 | 63 | 32 | 92 |
| 2 | 150 | G-CM+O-GMCM+ G-S+406080 | 65 | 34 | 90 |

We examined confusion matrices for vowels and for consonants to find the most confused vowels and consonants and with what they are confused with. We considered all the characters confused with 5 or more times. The vowel confusion matrix is as shown in the figure 11. For clarity the typed characters are shown instead of handwritten images.

| Sl No | Vowel and its matra | | Confused 5 or more times with another vowel and its matra | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | ಅ | | ಇ | ಿ | ಉ | ು | | |
| 2 | ಆ | ಾ | ಔ | ೌ | ಅಂ | ಂ | | |
| 3 | ಇ | ಿ | ಅ | | ಉ | ು | ಎ | ೆ |
| 4 | ಉ | ು | ಅ | | ಅಂ | ಂ | | |
| 5 | ಊ | ೂ | ಅಂ | ಂ | | | | |
| 6 | ಋ | ೃ | ಐ | ೈ | | | | |
| 7 | ಏ | ೇ | ಓ | ೋ | | | | |
| 8 | ಔ | ೌ | ಆ | ಾ | ಇ | ಿ | | |
| 9 | ಅಂ | ಂ | ಈ | ೀ | ಅಃ | ಃ | | |

Figure 10. Vowel Confusion Matrix

The analysis of the confusion matrix shows that the confusion is between the similar vowels matras. Also it is observed that some Kagunita shapes become similar to other Kagunita shapes.

For example, ಮ and ವು , ಲ and ತಿ, ತಾ and ಲಾ etc.

With huge writing variations, if the circle in (ಅಂ) ಂ matra becomes smaller or the circle in (ಆ) ಾ matra becomes bigger, there will be close similarities. If the (ಎ) ೆ matra is written with short horizontal line, it can be confused with (ಇ) ಿ

matra.

The Consonant confusion matrix is as shown in the figure 12. We can observe the close shape similarities as the reason for the confusion. For example ವ with ಉ ು matra becomes ವು with a shape very close to ಮ**.** The manual shape similarities observation matches with the confusion results. To reduce such confusions, we need to strengthen the feature set. This is part of our ongoing work.

| Sl No | Consonant | Confused 5 or more times with | | |
|---|---|---|---|---|
| 1 | ಕ | ಠ | | |
| 2 | ಘ | ಫ | | |
| 3 | ಚ | ಜ | | |
| 4 | ಜ | ಳ | | |
| 5 | ಝ | ಖ | | |
| 6 | ಡ | ದ | | |
| 7 | ಢ | ಥ | ಧ | |
| 8 | ತ | ಲ | | |
| 9 | ಥ | ಧ | ಢ | ಫ |
| 10 | ದ | ಡ | | |
| 11 | ಧ | ಢ | | |
| 12 | ಪ | ತ | | |
| 13 | ಫ | ಥ | ಘ | |
| 14 | ಬ | ವ | | |
| 15 | ಮ | ವ | | |
| 16 | ವ | ಮ | | |
| 17 | ಷ | ಫ | | |

Figure 11. Consonant confusion matrix

No comparable work in Indian language OCR exists for comparing our results; much less for Kannada. Much of the work covers simply numerals or at best stand alone vowels and consonants. Though, in absolute terms, the recognition figures may not be high enough for practical use; in a practical system one can use linguistic information to significantly enhance the effective recognition. For example, using a dictionary to cross check and rank the characters can often get to recognize words with high enough accuracy for practical use.

## VI. Conclusions

We experimented on moments as features for handwritten Kannada Kagunita recognition. From the above results, we find that the Kagunita recognition can be done by individual vowel and consonant recognition rather than as a Kagunita. This reduces the number of characters to be recognized from 510 to just 34 consonants and 15 vowels. That is only a total of 34+15=49 unique shapes need to be identified. We have found that the feature set most suited for vowels is different from those for consonants. Accordingly we have used two different classifiers, one for vowel classification and the other for consonant classification. We have compared the classification accuracy of moments features computed from original images with directional images. We find the moments from directional images as strong features. We also find a nonlinear relation in the recognition power of both the features and so can be used together as a single feature set for the recognition. The cut image concept worked well for the Kagunita recognition. The moments features from the cut images play an important role in the recognition of Kagunita. Under vowel recognition experiments, the three cuts from the right performed well over single cut on the right. Under consonant recognition experiments, the three cuts from the left performed better Also from the experiments it is clear that the cut size is not very sensitive within small variations. The confusion matrix analysis of both vowels and consonants shows that the existing confusion is between strongly similar shaped characters.

Hence the proposed technology of extracting moments features from directional images and from the cut images is promising and can be combined with other features to improve the performance further. We are now investigating an overall architecture for HCR incorporating all these, attempting to improve the overall accuracy further. Such a structure will help to exploit further domain information in the recognition process. For example, certain vowels modify mostly certain parts of the consonants. For practical use, as mentioned earlier, we can also use linguistic inputs.

## References

[1] Bhardwaj Anurag, Jose Damien and Govindaraju Venu. 2008. “Script Independent Word Spotting in Multilingual Documents.” In the Second International workshop on “ Cross Lingual Information Access-2008”.

[2] Chung Yuk Ying, Wong Man To and Bennamoun Mohammed, 1998, “ Handwritten Character Recognition by contour sequence moments and Neural Network”, In International Conference on Systems, Man, and Cybernetics, Volume 5, Issue , 11-14 Oct 1998 page(s): 4184 - 4188

[3] Hamamoto Y, Uchimura S, Masamizu K and Tomita S, 1995, “Recognition of Handprinted Chinese Characters using Gabor Features” In Proceedings of the 3rd International conference on Document Analysis and Recognition. Volume 2, page(s): 819-823.

[4] Huang Yu, Xie Mei. May 2005. “A Novel Character Recognition Method based on Gabor Transform.” In International Conference on Communications, Circuits and Systems. Vol 2. page(s): 815-819

[5] Kunte R Sanjeev and Samuel Sudhaker R D. December 2007. “A Bilingual Machine-Interface OCR for Printed Kannada and English Text Employing Wavelet Features.” In Tenth IEEE International conference on information technology. page(s): 202-207.

[6] Liao Simon X and Lu Qin, 1997, “A study of moment functions and its use in Chinese Character Recognition”, In Proceedings of the 4th International conference on Document Analysis and Recognition (ICDAR-97) page(s): 572-575

[7] Liao S X and Pawlak M. 1998. “On the Accuracy of Zernike Moments for Image Analysis.” In IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(12). page(s): 1358–1364.

[8] Liu C L, Koga M and Fujisawa H. 2005. “Gabor Feature Extractiion for Character Recognition: Comparison with Gradient Feature. “ In Eighth International on Document Analysis and Recognition. vol 1. page(s): 121-125.

[9] Mallat S G, 1998, “Multifrequency Channel Decompositions of Images and Wavelet Models,” In IEEE Transactions on Image Processing, 37(12), page(s): 2091–2110.

[10] Pal U; Chaudhuri B B, ‘Indian Script Character Recognition: A Survey’, In Elsevier Ltd. Pattern Recognition Society- 2004, page(s): 1887 – 1899,

[11] Philip Bindu and R.D. Sudhakar Samuel. “A Novel Binlingual OCR for Printed Malayam-English Text based on Gabor Features and Dominant Singular Values.” In IEEE International Conference on Digital Image Processing. page(s): 361-365.

[12] Ragha Leena and M Sasikumar. October 07. Fusion of Geometric, Positional and Moments features for Handwritten Character Recognition using Neural Network, In National Conference on Computer Vision, AI and Robotics – ( NCCVAIR - 07 ) at Chennai. India. page(s): 74-80.

[13] Ragha Leena and M Sasikumar. March 2009. Dynamic presprocessing and feature analysis for handwritten character recognition in Kannada, In IEEE International Advance Computing Conference (IACC-09) at Punjab, India. page(s): 3183-3188.
[14] Ragha Leena and M Sasikumar. "Using Moments Features from Gabor Directional Images for Kannada Handwriting Character Recognition" . In International Conference and Workshop on Emerging trends in Technology (ICWET-2010) at Mumbai, India. page(s): 125-130
[15] Ragha Leena and M Sasikumar. "Adapting Moments for Handwritten Kannada Kagunita Recognition" In International Conference on Machine Learning and Computing ( ICMLC-2010) at Bangalore. India. page(s): 125-129.
[16] Yamada Keiji. 1995. "Optimal Sampling Intervals for Gabor Features and Printed Japanese Character Recognition." Third International Conference on Document Analysis and Recognition. Vol 1, page(s): 150-153.
[17] Tian Xue-Dong, Zuo Li-Na, Yang Fang and Ha Ming-Hu, 2007, " An Improved method based on Gabor Feature for Mathematical Symbol Recognition", In 6th International Conference on Machine Learning and Cybernetics, Page(s): 1678-1682.

**Leena R Ragha** has done her BE in Computer Science and Engineering from GIT, Belgaum, Karnataka, India in 1991. She obtained her MS degree in Comuter Science and Engineering from BITS, Pilani, India in 1995. She is currently pursuing her PhD degree in Computer Science and Information Technology from SNDT university, Mumbai, India.
She is working with Ramrao Adik Institute of Technology an engineering institution under University of Mumbai since 1992. She is currently working as Assistant professor and was formerly head of Computer Engineering department. She is interested in research areas on Image Processing, Computer Vision and Pattern recognition.

**M Sasikumar** is currently Associate Director (Research) at CDAC, Kharghar, Navi Mumbai. He did his PhD (Computer Science) from BITS, Pilani; MS(Engg) from IISc Bangalore; and BTech(CS) from IIT Chennai.
He has been with CDAC Mumbai (formerly known as NCST) for over 20 years with focus on research and development. He has been involved in design and implementation of a number of projects - internal and external - in areas such as resource scheduling, natural language processing, data mining, e-learning, accessibility, healthcare, etc. Currently heads the R&D groups on artificial intelligence, educational technology, open source software, and language computing at CDAC Mumbai. Was chief investigator for the DIT funded project on national resource centre for online learning, and coordinator of the IBM-IIT-CDAC initiative for open source software resource centre. He is co-author of two textbooks - on expert systems, and on parallel computing and has anumber of publications under his credit.