CHAPTER 5

WRITER IDENTIFICATION BASED ON OFFLINE HANDWRITTEN DOCUMENT IN KANNADA LANGUAGE

5.1 Introduction

In earlier days it was a notion of the people who were using computers for their work have to adapt their style of input in a way computer expects—whether in typing, or filling out forms with letters. But at present, computers are able to accept different varieties of input methods such as human handwriting or biometric feature to name a few. At present computers are doing tasks which was once thought beyond their abilities. Pattern recognition systems are changing the way people relate to computers. This throws many challenges in the field of automatic pattern recognition:

• Is it possible to develop an algorithm that extracts the unique characteristics of each person?
• Which features are most important and how to represent them in the Computer?

Bio-metric systems are developed for this purpose. A typical bio-metric system consists of three parts; input device such as scanner or a writing pad, which are used to record the input for future processing. Biometric software which accepts the input in the digital form extracts the features and compares the results. A Database is used to store the features which will be used later for comparisons. These features extracted from the input samples saves space and time for processing.

Bio-metric system used for individual identification may be based on two types of Biological characteristics

i. Physiological  ii. Behavioral

Physiological system depends on the individual physical characteristics. Face recognition, Finger print authentication, Iris recognition falls in this category. Biological system based on the behavioral aspects relays on voice recognition, signature verification or handwriting of the individuals for identification.
Identification of a person based on Biometric features has gained much interest among researchers and has made transition from research level to full-fledged practical applications. Biometric features of each person are different and hence can be used effectively in distinguishing one person from another. Biometric system uses pattern recognition technology for identification and recognition of each individual’s. New applications are emerging in computer field such as Web searching, Data mining, Face recognition, Iris recognition and Handwritten recognition which requires robust and intelligent solutions. The solution to be developed requires recognition of a particular pattern and hence the field of pattern recognition is receiving more importance. One such Bio-metric feature is handwriting of a person. Identification of a person based on his/her hand written image is a challenging problem. The solution obtained can be effectively used in the field of historic document analysis, forensic etc.

 Millions of Handwritten documents are prepared each day and is a unique human act. Each writer can be characterized by his own handwriting. A few types of strokes are reproduced without his consciousness. As a result, handwriting of each individual plays a significant role as that of fingerprint in certain situations. Whenever there is a need to test the authenticity of a document, especially in the court of law or signature verification in Banks or similar other situations, writer identification system will be more useful.

**5.2 Some existing methods for writer identification**

Biometric features, which are used for personal identification, use the concept of pattern recognition [Plamond and G Lorette, 1989; A.K. Jain, R. Bolle, S. Pankanti, 1999; A.K. Jain, L. Hong, S. Pankanti, 2000; G.R. Ball and S.N. Srihari, 2009]. Biometric data of each individual is acquired, features are extracted and matched with the features available in a database to obtain most accurate results. These results may be used for variety of purpose including personal identification, crime or other security requirements. Many Biomedical features such as finger prints [D. Bayly, M. Castro, A. Arakala, J. Jeffers and K. Horadam, 2010] signature [C-Y. Low, A.B.J. Teoh and T. Connie, 2009; E. Marcu, 2010], facial features [T. Bourlai and J. Kittler, 2010; E. Norouzi, M.N. Ahmadabadi and B.N. Araabi, 2010] and other methods are used for this purpose.
Writer Identification systems are very popular and widely used biometric systems especially for the authentication of financial transaction. In many situations signature samples are not always available, especially in crime investigation, and also the quantity of data available for analysis in signature-based systems is relatively small. Hence writer identification systems have gained much importance in such situations. Handwriting samples have the advantage of being more easily available than signature samples and also provide more data for analysis. It is common practice to use handwriting for writer identification but it is done with the help of human experts in most cases.

Handwriting identification requires dissimilarity information. This helps to identify the writer. Figure 5.1 shows the standard frame work of writer identification. It begins with finding the writer among many writers in a large database with an assumption that each person’s has individualistic handwriting. Handwriting of each individual is influenced by his culture, style, type of the language and personalized differences to mention a few. This result in increased variations in shapes found in handwritten words; text lines may be curve lines and distance between the lines or between the words are non-uniform. Figure 5.2 shows handwriting samples of few writers and there exists a visible variation among them. This allows us to extract the unique features of each writer.

![Figure 5.1: Standard frame work for pattern recognition](image-url)
Writer identification based on their writing is one of the challenging areas in the field of pattern recognition and has a wide scope for research activities. Writings can be treated as a pattern and its recognition plays a vital role in human decision making [A.K. Jain, Fellow, IEEE, R.P.W. Duin, and J. Mao, 2000]. A writer identification system searches a database containing handwritings of known persons to find whether test handwriting matches or nearly matches with the existing one. Scientists are very keen to study the behavior of a person based on his writing using various tools including neurology, psychology, pattern recognition and data mining [Tan Gx, Viard-Gauden C, Kot A. C. 2009].

From the studies of various writer identification techniques it is observed that writer identification system falls into two category – text dependent and text independent. In the first approach better performance can be achieved with very small amount of writer data [Pavelec et al., zoop]. It uses comparison between individual character and words of known text content and requires segmentation of relevant information [Plamondon & Lorette 1989]. This has a demerit that it is applicable to a fixed text or human intervention is needed in extracting the objects of interest. Text independent approach even though has broad applicability requires large amounts of handwriting to obtain stable statistical features from the entire text block [Bularu M and Schomaker 2007]. Text dependent writer identification technique [Catalin I-Tomai, Bin Zhang and Sargur N Srihari 2003] works on the basis of writing style of a character or word by the writer. An offline text independent writer identification system for Arabic character [Mohamed Nidhal and Maher Khemakhem 2010] works on contour based features. Six different features, based on length, direction, angle and curvature measurements are used for this purpose.
Handwriting behavior of each individual can be identified from texture-level features. It gives information regarding habitual pen-grip and preferred style of each individual that depends on character shapes engrained in the motor memory and is influenced by educational, cultural and other factors on the writer [Imran Siddique and Nicole Vincent].

Writer identification is also of the type on-line or off-line [He et al 2008]. In online method the writing behavior of the writer is captured and converted into series of signals using on transducer device. In offline writer identification system, scanned image of the writer writing is used which depicts his behavior [Jin et al 2005]. Online writer identification system based on Temporal sequence code, which tracks speed and pressure change in writing, and shape codes that relay on direction of trajectory in writing was developed for Chinese and English language [Bangy Li and Tieniu Tan 2009]. It works better for small number of characters. Offline text independent writer identification using Hidden Markov Model [Andreas Schlapbach and Horst Bunke] works on the basis of computing the score unknown writer and comparing it scores of each individual writer. The score of each individual writer is computed by recognizer based on handwriting. The recognizer with the highest score is assigned as unknown writer. Online text independent writer recognition system [PitakThunswarin and TakenobuMatssura 2004] for Thai language based on velocity of Barycenter of pen-point movement using Fourier transform method. Another interesting work based on Shape code book was presented for Online Text-independent Writer identification, where each code word represents direction of writing trajectory, speed and pressure of writer [Bangy Li and Tieniu Tan].

In offline writer identification system, the hand written text of the writer is scanned and used for feature extraction. As such offline writer identifications poses more challenges compared to on-line method because of the absence of additional features, which are available to online systems, is absent for offline system. Writer identification for non-uniformly skewed handwriting images and statistical based method has been discussed in [H.E.S Said, G.S. Peake, T. N. Tan, and K. D. Baker] [Bulacu, M., Schomaker, L. and Vuurpijl, L, 2003]. Other methods of writer identification based on Hierarchical Shape Primitive features, Contour based features has been discussed in [B.Li, Z.Sun, and T.N.Tan][Mohamed NidhalAbdi and Maher Khemakhem].
Writer identification systems have been developed for languages like English, Arabic, Chinese and Indian languages such as Hindi, Bangla. Many features of the writing such as style, loops, distance between the characters, skewness of the line etc., forms the basic characteristic of a writer and can be used to identify the writer.

Inspired by the challenges and need of such a system for identification of writer in Kannada language, we have developed a technique for writer identification. The features of each individual writer, based on text image of the handwriting, are extracted. These features are stored and then used for identification of the writer. Empirical Mode Decomposition (EMD) is used to find the features of writers by computing the intrinsic energy and residue energy of each writer. EMD is basically designed for processing of non-liner non-stationary data [N.E.Huang, Z.Shen, S.R.Long]. However, it can be used to decompose the complex data into a finite and often a small number of Intrinsic Mode Functions (IMFs). This method is adaptive in nature and has better efficiency. It has been found, this method can be used for extracting the features in an image and other fields of science and engineering [Norden E. Huang and Zhaohua Wu]. Using this method, distinguishing features of each individual writer are effectively extracted and are stored in a database. Among several IMFs, which represents the writer features with different energy, first four IMFs along with energy of residue component is used for writer identification.

The organization of the remaining chapter is as follows: Section 5.3 gives the outline of the method proposed. Details of the algorithm used for extracting the feature of each writer are presented in the section 5.4. Experiment results are discussed in the section 5.5 and chapter is concluded in section5.7

5.3 Outline of the proposed method

A Design of a better pattern recognition system requires a standard database which will help in evaluating the algorithms and comparison of different methods. This, however, is not available for most of the Indian languages including Kannada. To check the usefulness of our method for extracting the features of the individual writers, we created our own database. Writer samples are collected, scanned and stored as image files. From each writer, 5 samples of handwriting is collected at different time of the day to take care of possible variation in
their writing due to change in their psychological behavior at different time of the day. This may due to stress level or their psychological attitude.

In order to identify the writer, it is necessary to extract characteristic feature of each writer. The samples of the writer handwriting are scanned to convert it into digital form. These scanned images cannot be directly used as it may contain noise. A dusty scanner or scratches on the scanner’s glass or paper on which handwritten sample is available may be mutilated which introduces noise into scanned image. Hence the image is preprocessed to remove the salt and pepper noise. Morphological operations – dilation and erosion are carried out to remove the noise from the image. Dilation followed by erosion removes the salt noise – referred to as closing operation and erosion followed by dilation removes pepper noise called as opening operation.

To extract the features, the image is divided into uniformly sized blocks. This is with a view that a part is a representative of a whole-text and the time required for features extraction will be greatly reduced. Input samples take more space on secondary memory as compared to mathematical data. To reduce the space the features of the writer handwriting which was obtained at different time of the day is computed and averaged. This process is carried out for all writers and features of different writers are stored. These features computed once is more useful and saves a lot of processing time.

The features obtained are then classified. For better recognition rate, multiple samples for each individual are collected during registration process. After classification, average value of the feature is computed for each writer. In brief, the method adopted is stated in the below steps:

Step 1: Take a sample of the writer handwriting.
Step 2: For each sample repeat steps 3 to 5.
Step 3: Find the features for each writer.
Step 4: Compute the average value of features selected in step 3.
Step 5: Compute the average value of each member in the group.
Step 6: Use the computed value in step 5 to get overall classification of the writer handwriting.
5.4 Feature extraction of each writer

Empirical Mode Decomposition developed method Huang et al. for the study of nonlinear and non-stationary properties of a time series. This method contains the following two steps: Empirical Mode Decomposition (EMD) and Hilbert Spectra Analysis (HSA). The main idea of EMD is to locally estimate a signal as a sum of a local trend and a local detail: the local trend is a low frequency part, and the local detail a high frequency. When this is done for all the oscillations composing a signal, the high frequency part is called an Intrinsic Mode Function (IMF) and the low frequency part is called the residual. The procedure is further applied again to the residual part from previous process, considered as a new times series, new IMF and a new residual extracted. This method of decomposition process, expresses a time series $x(t)$ as the sum of a finite number of IMFs $C_i(t)$ and a final residual $r_n(t)$.

5.4.1 Extraction of the Intrensic mode signal

An IMF two conditions: (i) the difference between the number of local extrema and the number of zero-crossings must be zero or one; (ii) the running mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. The following steps extract the IMF from the image.

Step 1: The local extrema of the image $x(t)$ are identified;
Step 2: The local maxima are connected together forming an upper envelope $e_{\text{max}}(t)$, which is obtained by a cubic spline interpolation. The same is done for local minima, providing a Lower envelope $e_{\text{min}}(t)$;
Step 3: The mean is defined as $m_i(t) = (e_{\text{max}}(t) + e_{\text{min}}(t))/2$;
Step 4: The mean is subtracted from the signal, providing the local detail $h_i(t) = x(t) - m_i(t)$;
Step 5: The component $h_i(t)$ is then examined to check if it satisfies the conditions to be an IMF. If yes, it is considered as the first IMF and denoted $C_i(t) = h_i(t)$. It is subtracted from the original signal and the first residual, $r_i(t) = x(t) - C_i(t)$ is taken as the new series in step.
5.4.2 Sifting process

If \( h_1(t) \) is not an IMF, a procedure called “sifting process” is applied as many times as needed to obtain an IMF. In the sifting process, \( h_1(t) \) is considered as the new data; the local extrema are estimated, lower and upper envelopes are formed and their mean is denoted \( m_1(t) \). This mean is subtracted from \( h_1(t) \), providing \( h_{11}(t) = h_1(t) - m_1(t) \). Then it is checked if \( h_{11}(t) \) is an IMF. If not, the sifting process is repeated, until the component \( h_{1k}(t) \) satisfies the IMF conditions. Then the first IMF is \( C_1(t) = h_{1k}(t) \) and the residual \( r_1(t) = x(t) - C_1(t) \) is taken as the new series in step 1.

By construction, the number of extrema decreases when going from one residual to the next; the above algorithm ends when the residual has only one extremum, or is a constant, and in this case no more IMF can be extracted. The complete decomposition is then achieved in a finite number of steps, of the order \( n \leq \log_2 N \), for \( N \) data points. The signal \( x(t) \) is can be written as:

\[
x(t) = \sum_{i=1}^{N} C_i(t) + r_i(t)
\]

(5.1)

EMD is a time-frequency analysis since it can represent the original signal in a energy-frequency-time form.

In the second step Hilbert transform is applied to each IMF. It is possible to design the Hilbert spectrum \( H(\omega, t) \), which represent the energy as the function of instantaneous frequency and time. Here the Hilbert transform is a singular integration, it can be taken as the best local fit of an amplitude and phase varying trigonometric function to \( x(t) \). Therefore the Hilbert spectrum can provide sufficient locality information. The locality and adaptivity abilities make this method unique and suitable for nonlinear and non-stationary time series analysis. Hence the concept of HHT can be conveniently used for extracting the unique character of each writer. The average energy of each writer can be computed as

\[
e_s = \frac{1}{2N} \sum_{i=1}^{N} [A_s^1(i) + A_s^2(i)] \text{Where}
\]

(5.2)

\[
A_s^j(i) = \sqrt{[\text{imf}_s^j(i)]^2 + [H(\text{imf}_s^j(i))]^2} \text{ for } j=1,4
\]

(5.3)
Where \( N \) is number of and \( H(imf_i^j(i)) \) is the Hilbert transform of \( imf_i^j(i) \).

Average energy of the residue for each writer is given by

\[
r^*_i = \frac{1}{N} \sum_{i=1}^{N} res_i(i)
\]  

(5.4)

### 5.5 Writer Identification

Once the features of each writer is obtained and stored, these features are then can be used as input to the writer classifier. The common issues found in identification system commonly involved preserving dissimilarity features of each writer, finding the best discriminating features, selecting a method to compare them and classify them that identifies the writer. To find the closest matching of the features of unknown writer with the features of the known writer we have used nearest neighbor classification method. This works on the basis of ‘leave one out’ strategy. In this rule, distance between two writer images is an important criterion to be considered. The advantage of NN classifier is, without a priori about the distributions from which training samples are drawn, this rule achieves better performance. If the writer in the sample image is identical or near to the images in the class then distance measurement will be lower else the image does not belongs to that class. After measuring the distance of the image, writer identity will be assigned to the most similar class that the writer belongs. The following equation gives the details of the comparison between the known writer and unknown author.

\[
d_k = \sqrt{\left[ \sum_{j=1}^{N} (u_j - f_{k,j})^2 \right]}
\]  

(5.6)

d_k: class distance of unknown writer

\( u \) : Feature of unknown writer

\( f_{k,j} \) : best feature vector, \( j \) of class \( k \)

\( j \) : \( 1 \ldots n \) number of members in that class

\( k \) : \( 1 \ldots \ldots \) Number of writers
If a match exists, then $d_j$ for that writer $k$ is small compared to all other writers and unknown writer is identified as writer $k$.

5.6 Experimental results

To check the suitability of the proposed method, we conducted experiments on the samples of the writings of the different writers. The results obtained are stored and is compared with feature extracted from unknown author to decide the authenticity of unknown author.

5.6.1 Extraction of writer features

The unique features corresponding to each writer is extracted by knowing that each writer has his/her own style of writing. Writer’s handwriting is scanned using HP Scanner with a resolution of 200DPI. For each writer samples of handwriting is taken at different time of the day. The scanned image is processed to remove the noise. To avoid the influence of blank spaces in the text line, text padding is done. An image of size 100x100, as shown in the Figure 5.3, is used for the experiment purpose. For the experiment purpose, we considered 5 samples of 30 writers along with machine printed one. We use four sets of handwritings to mark the feature of the writer. One set is used for testing purpose. The feature obtained is then averaged. For each block of samples, IMFs and residue component is found. Since EMD acts as a filter, the information associated with each IMF keeps decreasing; finally, a stage is reached where further processing is not possible. The remaining signal is called residue. To obtain the feature of each writer, first four IMFs along with the residue are considered. The first four IMFs for two different writers are shown in Figure 5.4. The energy associated with them is computed. Average high frequency energy and average residue energy is computed. These energy components act as a feature for each writer and are stored for future use.
Figure. 5.3: Writing Sample of Writers used for feature extraction

Figure. 5.4: The first four IMFs of the two writers
A number of experiments have been conducted on the writings of 50 people. Features extracted from each writer is stored, Identification of the new writer is performed using k-nearest neighbor classifier. The results achieved are promising and an identification accuracy of 68% to 94% was obtained (100% for machine printer character). The Table 5.1 and 5.2 shows the high frequency energy and low frequency energy of different writers. Table 5.3 gives the recognition rate achieved for different writers.

**Table 5.1: High frequency energy of different writers**

<table>
<thead>
<tr>
<th>Writer</th>
<th>Writing1</th>
<th>Writing2</th>
<th>Writing3</th>
<th>Writing4</th>
<th>Writing5</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writer1</td>
<td>.1846</td>
<td>.1874</td>
<td>.1897</td>
<td>.1901</td>
<td>.1830</td>
<td>0.1869</td>
</tr>
<tr>
<td>Writer2</td>
<td>.1506</td>
<td>.1496</td>
<td>.1522</td>
<td>.1531</td>
<td>.1489</td>
<td>0.1508</td>
</tr>
<tr>
<td>Writer3</td>
<td>.1668</td>
<td>.1677</td>
<td>.1698</td>
<td>.1701</td>
<td>.1634</td>
<td>0.1675</td>
</tr>
<tr>
<td>Writer4</td>
<td>.1367</td>
<td>.1420</td>
<td>.1409</td>
<td>.1388</td>
<td>.1428</td>
<td>0.1402</td>
</tr>
<tr>
<td>Writer5</td>
<td>.0536</td>
<td>.0610</td>
<td>.0586</td>
<td>.0519</td>
<td>.0525</td>
<td>0.0555</td>
</tr>
<tr>
<td>Writer6</td>
<td>.0702</td>
<td>.0792</td>
<td>.0715</td>
<td>.0808</td>
<td>.0719</td>
<td>0.0747</td>
</tr>
<tr>
<td>Writer7</td>
<td>.0485</td>
<td>.0435</td>
<td>.0402</td>
<td>.0502</td>
<td>.0488</td>
<td>0.0462</td>
</tr>
<tr>
<td>Writer8</td>
<td>.0469</td>
<td>.0480</td>
<td>.0429</td>
<td>.0463</td>
<td>.0519</td>
<td>0.0472</td>
</tr>
<tr>
<td>Writer9</td>
<td>.2430</td>
<td>.2476</td>
<td>.2398</td>
<td>.2447</td>
<td>.2492</td>
<td>0.2449</td>
</tr>
<tr>
<td>Writer10</td>
<td>.1920</td>
<td>.1956</td>
<td>.1897</td>
<td>.1902</td>
<td>.1938</td>
<td>0.1922</td>
</tr>
</tbody>
</table>
**Table 5.2:** Low frequency energy of different writers

<table>
<thead>
<tr>
<th>Writer</th>
<th>Writing1</th>
<th>Writing2</th>
<th>Writing3</th>
<th>Writing4</th>
<th>Writing5</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writer1</td>
<td>.6365</td>
<td>.6875</td>
<td>.6672</td>
<td>.6491</td>
<td>.6479</td>
<td>0.6576</td>
</tr>
<tr>
<td>Writer2</td>
<td>1.1166</td>
<td>1.0897</td>
<td>1.2132</td>
<td>1.1284</td>
<td>1.1945</td>
<td>1.1485</td>
</tr>
<tr>
<td>Writer3</td>
<td>.7250</td>
<td>.7177</td>
<td>.7195</td>
<td>.7206</td>
<td>.7269</td>
<td>0.7219</td>
</tr>
<tr>
<td>Writer4</td>
<td>.4729</td>
<td>.4763</td>
<td>.4693</td>
<td>.4701</td>
<td>.4782</td>
<td>0.4734</td>
</tr>
<tr>
<td>Writer5</td>
<td>1.1207</td>
<td>1.1324</td>
<td>1.2007</td>
<td>1.0989</td>
<td>1.1340</td>
<td>1.1373</td>
</tr>
<tr>
<td>Writer6</td>
<td>.3916</td>
<td>.4012</td>
<td>.4140</td>
<td>.3902</td>
<td>.4076</td>
<td>0.4009</td>
</tr>
<tr>
<td>Writer7</td>
<td>1.3166</td>
<td>1.3086</td>
<td>1.3122</td>
<td>1.3212</td>
<td>1.2988</td>
<td>1.3115</td>
</tr>
<tr>
<td>Writer8</td>
<td>1.1316</td>
<td>1.1385</td>
<td>1.1293</td>
<td>1.1304</td>
<td>1.1346</td>
<td>1.1329</td>
</tr>
<tr>
<td>Writer9</td>
<td>1.2582</td>
<td>1.2597</td>
<td>1.2613</td>
<td>1.1940</td>
<td>1.2302</td>
<td>1.2407</td>
</tr>
<tr>
<td>Writer10</td>
<td>.9634</td>
<td>.9547</td>
<td>.9601</td>
<td>.9594</td>
<td>.9619</td>
<td>0.9599</td>
</tr>
</tbody>
</table>

**Table 5.3:** Recognition rate of different users

<table>
<thead>
<tr>
<th>Writer1</th>
<th>Writer2</th>
<th>Writer3</th>
<th>Writer4</th>
<th>Writer5</th>
<th>Writer6</th>
<th>Writer7</th>
<th>Writer8</th>
</tr>
</thead>
<tbody>
<tr>
<td>72.3</td>
<td>76.4</td>
<td>89.3</td>
<td>90.2</td>
<td>94.1</td>
<td>82.1</td>
<td>87.3</td>
<td>85.4</td>
</tr>
<tr>
<td>68.3</td>
<td>69.5</td>
<td>87.6</td>
<td>84.6</td>
<td>82.9</td>
<td>81.3</td>
<td>89.8</td>
<td>91.5</td>
</tr>
<tr>
<td>Writer17</td>
<td>Writer18</td>
<td>Writer19</td>
<td>Writer20</td>
<td>Writer21</td>
<td>Writer22</td>
<td>Writer23</td>
<td>Writer24</td>
</tr>
<tr>
<td>90.3</td>
<td>69.6</td>
<td>69.9</td>
<td>77.3</td>
<td>83.6</td>
<td>85.7</td>
<td>82.1</td>
<td>74.9</td>
</tr>
<tr>
<td>Writer25</td>
<td>Writer26</td>
<td>Writer27</td>
<td>Writer28</td>
<td>Writer29</td>
<td>Writer30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>88.2</td>
<td>91.7</td>
<td>87.6</td>
<td>68.6</td>
<td>84.9</td>
<td>86.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5.7 Conclusion

In this chapter, a novel method for writer identification is presented. The need for such a system and its practical importance has been discussed. The writings in Kannada language are considered for this purpose. Each writer has a unique style of writing which distinguishes him from others. These features of each writer are computed from their handwritten document images. Empirical mode decomposition is a powerful tool which can be used to obtain the feature of each individual writer. The result obtained by this approach can be used for personal identification. This method is global in nature, as such there is no need to perform line and word segmentation thereby considerable amount of processing time is saved. The features of each writers handwriting is in the form of intrinsic mode function and a residue component. This represents unique characteristics of each writer. This feature of each writer is used for identification.

The feature computed is influenced by the many factors. Each writer has his/her own preference regarding pen type, color of the ink used, thickness of the handwriting etc. It is assumed that these habits are unique to each writer. Each writer most of the time prefers to use particular type of pen, ink color, and a few prefers to write with thick strokes and few others with thin strokes. A few writers have the habit of writing tilted downwards or upwards or some times wavery in nature. These factors also contribute in identification of the writer. The IMF derived is influenced by these parameters. The merit of this method is test sample size and training sample size may be different. Another advantage of this method is it works on the gray scale images and hence requires no binarization.

Additional characteristics of each writer and advanced classifiers may improve the accuracy of writer identification. Investigations are in progress for further improvement of our approach.