

## Inheritance of Handwriting Features

Monika Saini<sup>1</sup>, Aishwarya Srinivasan<sup>2</sup> and Anup Kumar Kapoor<sup>3</sup>

<sup>1</sup>UGC-JRF and <sup>3</sup>Professor, Department of Anthropology, University of Delhi, Delhi-110007,

<sup>2</sup>Engineer, Vellore Institute of Technology, Chennai-600048

**Abstract:-** Handwriting is one of the acquired characteristics of humans. It is a mixture of nature and nurture. Parents play an important part in teaching pre-writing skills to their children. Genetics also has a role in shaping the writing habits of the writer (such as handedness and handwriting positions). This paper examines the handwritings of parents and their biological off-springs to determine the inheritance of handwriting features from one generation to the next generation. In this scheme, resemblances in the handwriting features of parents-off springs were studied using computational features based on MATLAB 8.3 software.

**Keywords:-** Genetics, Inheritance, Resemblance, Handwriting Features.

### I. INTRODUCTION

Handwriting is a neuromuscular task which depends on the cognitive skill and a coordinated hand-eye movement of the writer. The development and progress of one's handwriting passes through four stages in the course of a lifetime. (1) the formative stage, (2) the impressionable or adolescent stage, (3) the mature stage, (4) the stage of degeneration. It is the learning stage for all persons, where letter designs are being developed and writing skills are being mastered. In the latter stages, changes are likely to occur in the fluency or design; the former usually being the result of writing frequency and, the latter usually being the product of a particular occupation.

Parents are the children's learning models. A preschooler is influenced by the handwriting of his or her parents. He learns to write letters by imitating or copying the letters formed by the parents. For example, a son may like the shape of a letter written by his father so he practices making that letter until he thinks it is like his father's. When he is satisfied with the way he writes it, how it looks (because it is like his father), he will begin to write it without paying attention to its details. By repeated attempts to copy this basic pattern the child incorporates the peculiarities of writing of his parents into his writing habits. The better his memory of forms and the more time he spends practicing letter forms written by his parents, the more habitual his movement becomes and the closer the new writing form will be to the parental writing model.

This tendency also occurs during adolescence, when the writer is developing and experimenting with his/her writing. As a writer matures he incorporates more individuality in his writing but he is still influenced by the previous writing system features and his writing may very well become a combination of the two. How much of a combination depends upon the writing habits he develops (Morris, 2000).

Writing habits of the off-springs are also shaped by genetic factors. These writing habits may include handedness, writing posture, pen position, learning and thinking ability. A few studies have also been done on familial resemblance with respect to hand preference, handwriting posture and specific cognitive abilities. (McKeever, 1979 & 2000; DeFries et al, 1979)

### II. RELATED WORK

The earliest published comment respecting the possibility of writing similarities due to familial relationship is one in the December 2, 1911 issue of *Scientific American* that reports an article by R.H. Chandler. The article illustrates three cases in which some general similarity is noted between brothers, sisters or fathers and sons, but the similarity was not so great. Stevens () conducted a comparison of the writings from three control groups in with the writings of eight members (seven females) of one family, all of whom were dextral writers, taught in the same schools by the same teachers and with the palmer writing system. The control groups were matched as closely as possible in age and sex, but were unrelated to each other. Nine elements were said to have been considered including stroke endings, continuity, slant, letter forms, proportions, size, speed, skills and pictorial appearance. The results indicated that there was greater similarity and less variation between the members of the family than there was in the control groups. Munch (1987) reported on the similarity found in one case in the writings of a mother and daughter. The questioned material was limited to three sets of initials. While there were disparities between the questioned and known writings they were not so consistent or so fundamental that they could not have been attributed to natural variations.

The present study concerns with similarities in the handwritings of parents and their off-springs using computational method. The writers also wished to determine, first whether like sex siblings show a greater

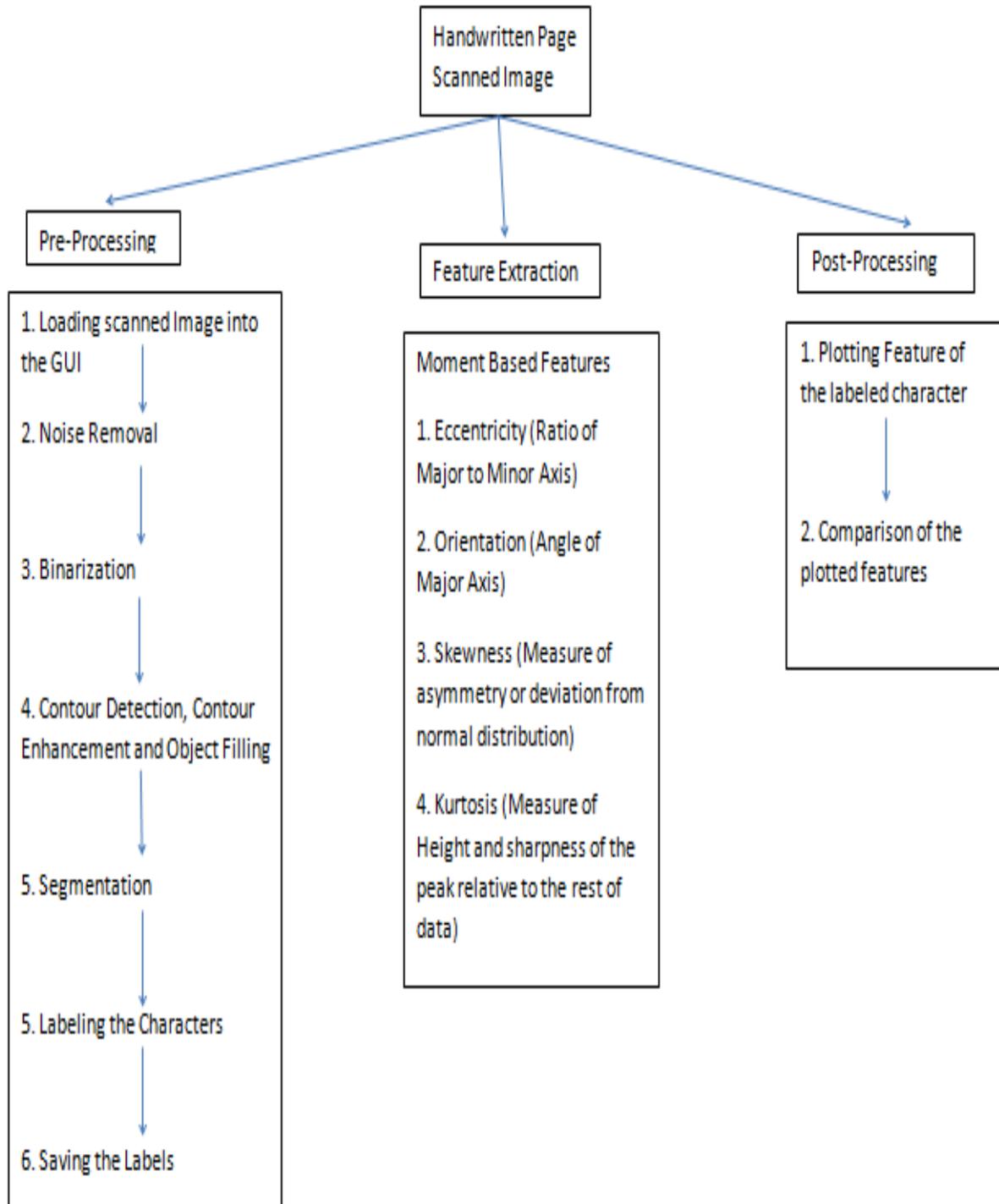
degree of resemblance than unlike sex; and second whether older off-springs show a greater degree of resemblance than younger off-spring.

In the next section we describe about three processing stages: Pre-Processing, Feature Extraction and Post-Processing. Section III is concerned with experimental tests and results. Finally some conclusions are drawn in section 4.

### III. METHODOLOGY

Similar to any pattern recognition system, the handwritten recognition system in MATLAB contains the following basic components: Pre-Processing, Feature Extraction and Post-Processing.

The basic structure for the proposed study is as follows:



**Figure 1: Basic Structure of the Proposed System**

### 1. Pre-Processing

This section elaborates the details of the pre-processing shown in Figure 1:

#### *Loading scanned image:*

Scanned Handwritten image is imported into the Graphical User Interface by clicking on the image menu and selecting open. Both JPG and TIF file formats are supported.

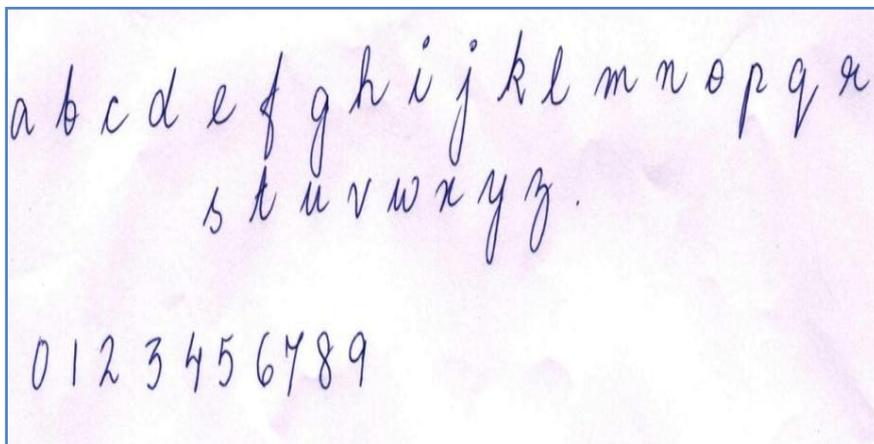
#### *Noise Removal (Median Filtering)*

Median filtering is a nonlinear method used to remove noise from images. It is widely used as it is very effective at removing noise while preserving edges. The median filter works by moving through the image pixel by pixel, replacing each value with the median value of neighboring pixels. The pattern of neighbours is called the "window", which slides, pixel by pixel over the entire image pixel, over the entire image. The median is calculated by first sorting all the pixel values from the window into numerical order, and then replacing the pixel being considered with the middle (median) pixel value.

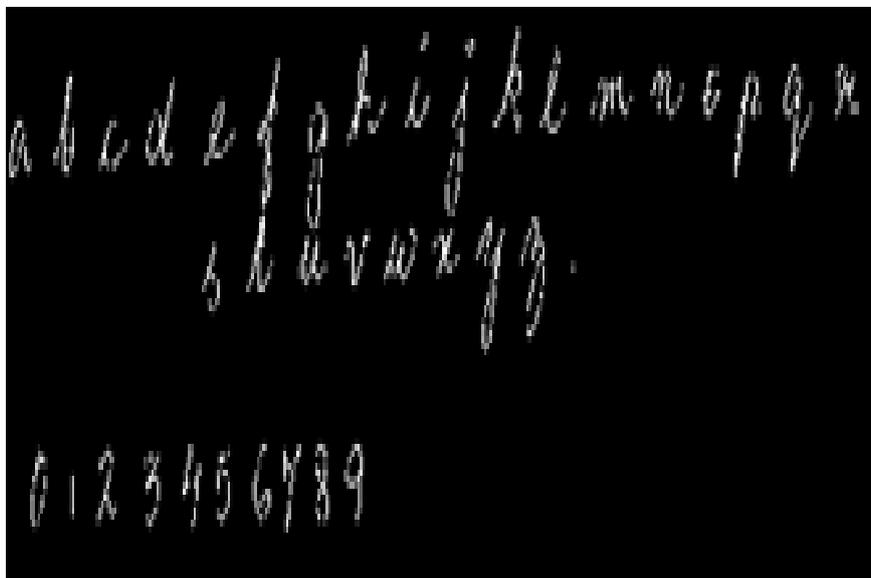
#### *Binarization*

Binarization or thresholding is such a widely-used method and generally, its process is to first determine a gray threshold according to some objective criteria and then assigns each pixel to one class (such as the foreground) if its gray level or gray value is greater than the determined threshold and otherwise to the other class (such as the background) (Chen & li, 2006).

If the input image is not a grayscale image, MATLAB converts the input image to grayscale, and then converts this grayscale image to binary image by thresholding.



Scanned Image



Binary Image

***Contour Detection, Contour Enhancement and Object Filling***

Contour detection or Edge detection is an image processing technique for finding the boundaries of objects within images. It works by detecting discontinuities in brightness. It is used for image segmentation and data extraction in areas such as image processing, computer vision, and machine vision.

Contour Enhancement or Image Dilation adds pixels to the boundaries of objects in an image. The value of the output pixel is the maximum value of all the pixels in the input pixel's neighborhood. In a binary image, if any of the pixels is set to the value 1, the output pixel is set to 1.

Object filling or Image filling is the process of filling holes in a binary image. A hole is a set of background pixels that cannot be reached by filling in the background from the edge of the image.

***Segmentation***

Segmentation refers to the process in which an image is subdivided into constituent regions or objects. These objects can be further processed or analyzed for the extraction of quantitative information.

***Labeling the Segmented Character***

After segmentation process, each segmented character will appear in the box below the binarization and segmentation command box. With the help of class label command, a user can label the character.

***Saving the Labels***

Segmented character and labels can be saved by clicking on the data menu and then selecting save.

**2. Feature Extraction**

Each character image samples can be expressed as set of quantifiable attributes called features. Feature Extraction is defined as the problem of "extracting from the raw data the information which is most relevant for classification purposes, in the sense of minimizing the within-class pattern variability while enhancing the between-class pattern variability (Devijver & Kittler, 1982). Features can be classified into two main categories, structural features (like strokes end points, intersections of line segments etc.) and statistical features (derived from statistical distribution of points like zoning, moments) (Heutte et al, 1998). In statistical approach, a character image is represented using a set of n features which can be considered as a point in n-dimensional feature space.

Shape is an important visual feature and it is one of the basic features used to describe image content. Shape representation and description techniques can be generally classified into two class of methods: contour-based methods and region-based methods. The classification is based on whether shape features are extracted from the contour only or are extracted from the whole shape region. In the region based method, moment descriptors are very useful tools to describe shapes. Moments can provide characteristics of an object that uniquely represent its shape. Important information about a shape such as its size, center location and orientation are all moment-based attributes. The 2-D moments of the character are:

$$m_{pq} = \int \int xy^q f(x,y) \, dx \, dy \quad p,q=0,1,2L$$

Moment's  $x$   $y$  theory can be used to analyze the contour and region of an object. In the present study, statistical and region based features are taken into consideration. The region based descriptors include moment invariants, Zernike moments, Radial Chebyshev moments, etc.

Regional feature vectors like Centroid, Eccentricity, Orientation, Skewness, Kurtosis, Mean and Standard deviations are also calculated.

Centroid- It is the measure of the center of the mass of the region.

Eccentricity- It is the scalar that specifies that eccentricity of the ellipse that has the same second moments as the region.

Orientation- It is the scalar value which is the measure of angle in degrees between the x-axis and the major axis of the ellipse that has the same second moments.

Skewness- It is a measure of the degree of asymmetry of the distribution of a variable about its mean.

Kurtosis- It is the measure of the height and sharpness of the peak relative to the rest of the data.

Mean- It is defined as the average value of a region.

Standard deviation- It is the measure of the dispersion of a set of data from its mean.

**IV. EXPERIMENTAL TEST AND RESULT**

For the experimental purpose, a data set (Handwriting sample) of 130 families with an average of 3 members in each family (Two Parents and one/two offspring) was obtained. The samples were then scanned

through high resolution scanner at 600 dpi and then it was used for further analysis. The similarities in handwriting features for familial resemblance were expressed by percentage similarity.

**Table 1: Average Percentage Similarity of Handwriting Features among Closed Genotypic Family Members**

| Relationship             | Average Percentage Difference (%) |             |          |          |            |
|--------------------------|-----------------------------------|-------------|----------|----------|------------|
|                          | Eccentricity                      | Orientation | Skewness | Kurtosis | Cumulative |
| <b>Father – son</b>      | 99.08                             | 94.94       | 98.48    | 99.00    | 97.87      |
| <b>Father – daughter</b> | 99.27                             | 94.90       | 98.07    | 98.99    | 97.80      |
| <b>Mother – son</b>      | 98.43                             | 95.11       | 98.21    | 98.96    | 97.67      |
| <b>Mother – daughter</b> | 98.23                             | 94.15       | 97.73    | 98.92    | 97.25      |
| <b>Siblings</b>          | 98.16                             | 93.87       | 97.01    | 98.79    | 96.95      |

It is seen from the result that the orientation feature of the handwriting has the least similarity among different genotypic relationships the least inherited characteristics. On the other side, father-son relationship shows maximum percentage of similarity whereas siblings reflects minimum percentage of resemblances in handwriting features.

## V. CONCLUSION

The experiment was conducted to find the influence of inheritance on handwriting characteristics. The similarities in the behavior and anatomy of a person are dependent on the inherited characters which in another way influence the handwriting. In the present study the father-son relationship has most of the similarities in handwriting features which may be due to the same sex influence and male gene dominance. Similarities obtained in handwriting features of parent-off spring and siblings validates that genes play an important role in transmission of writing habits from one generation to next generation and within the same generation.

## ACKNOWLEDGEMENTS

The authors are grateful to all the subjects for their cooperation in data collection. Ms. Monika Saini is equally thankful to University Grant Commission for providing financial support in the form of Junior Research fellowship. The authors also acknowledge Mathworks for developing Matlab software, which immensely benefitted in completion of this research.

## REFERENCES

- [1]. Chen, S., & Li, D. (2006). Image binarization focusing on objects. *Neurocomputing*, 69(16), 2411-2415.
- [2]. DeFries, J. C., Johnson, R. C., Kuse, A. R., McClearn, G. E., Polovina, J., Vandenberg, S. G., & Wilson, J. R. (1979). Familial resemblance for specific cognitive abilities. *Behavior Genetics*, 9(1), 23-43.
- [3]. Devijver, P. A., & Kittler, J. (1982). *Pattern recognition: A statistical approach*(Vol. 761). London: Prentice-Hall.
- [4]. Hansen, J. (2002). A Matlab Project in Optical Character Recognition (OCR). *DSP Lab, University of Rhode Island*, 6.
- [5]. Heutte, L., Paquet, T., Moreau, J. V., Lecourtier, Y., & Olivier, C. (1998). A structural/statistical feature based vector for handwritten character recognition. *Pattern recognition letters*, 19(7), 629-641.
- [6]. Hu, M. K. (1962). Visual pattern recognition by moment invariants. *Information Theory, IRE Transactions on*, 8(2), 179-187.
- [7]. McKeever, W. F. (1979). Handwriting posture in left-handers: Sex, familial sinistrality and language laterality correlates. *Neuropsychologia*, 17(5), 429-444.
- [8]. McKeever, W. F. (2000). A new family handedness sample with findings consistent with X-linked transmission. *British Journal of Psychology*, 91(1), 21-39.
- [9]. Morris, R., & Morris, R. N. (2000). *Forensic handwriting identification: fundamental concepts and principles*. Academic press.