Human Behavioral Analysis Based on Handwriting Recognition and Text Processing

Poorna Banerjee Dasgupta

M.Tech Computer Science and Engineering, Nirma Institute of Technology
Ahmedabad, Gujarat, India

Abstract

Behavioral characteristics are one of the primary features that help in differentiating one person from another. Numerous studies have been conducted that have analyzed the psychological and emotional aspects of human behavior for a multitude of applications, ranging from anthropological studies and medical diagnoses to solving cases of crime. However, it is a lesser known fact that human behavior and psychology can be adjudged by analyzing an individual's written texts and handwriting. Hence, this paper proposes and discusses an algorithmic approach for analyzing human behavioral characteristics through text processing and handwriting recognition, with the objective of incorporating the obtained results with futuristic artificial intelligence systems that can employ text processing and handwriting recognition as individualistic signatory features.

Keywords— *artificial intelligence, handwriting recognition, human psychology, text processing.*

I. INTRODUCTION

Comprehending human behavior and psychology in its entirety is an extremely complex task that integrates several fields of research. Consequently, several studies have been conducted that have analyzed the psychological and emotional aspects of human behavior for a myriad of applications, such as anthropological studies, medical diagnoses, forensic sciences, and solving cases of crime. Humans are known to exhibit individualistic behavioral characteristics in several forms such as speech, postures and gestures, and facial expressions [1, 2].

An equally important and unique aspect through which humans display their psychology is written texts. The analysis of one's text, whether handwritten or digital, can surprisingly reveal several facets about an individual's personality. It is a well-known fact that the handwriting of a human can vary from time to time, depending on one's emotional and psychological state [3]. Several learning disabilities in children such as dyslexia, dysgraphia, and ADHD are often diagnosed by the affected child's inability to maintain coherence and consistency while writing similar looking characters, or finding difficulty in spelling words and expressing thoughts on paper [4].

There are various other instances where the medium of writing has been employed as a tool to exhibit and affect human psychology. For example, famous authors and poets such as William Shakespeare, William Wordsworth, and Samuel Taylor Coleridge (to name a few) have all had their own distinct styles of writing. Commercial firms and magazine editorials put in a lot of thought and effort while creating brand advertisement mottos and writing articles so as to influence greater public masses. Forensic science experts often rely on text processing and handwriting recognition while solving cases of homicide that involve suicide notes written by the victim, or while trying to identify the sociopathic tendencies of abductors and criminals who leave behind ransom notes.

The aforementioned instances amply demonstrate that text processing and handwriting recognition play a crucial role in diagnosing and analyzing human behavioral characteristics. *Graphology* is a projective technique that is employed for analyzing handwriting based on the study and description of strokes of handwriting, which can reveal personality traits of the writer [5, 6]. The graphical features of handwriting are examined in terms of variations in the movement, shape, and space of the graphical writing gesture. There are primarily eight graphical features that can be analyzed through graphology: *layout, dimension, connection, form, slant, baseline, speed, and pressure* [7].

Layout or order is the overall impression provided by the written text. It is related to the degree of clarity in the ideas of the writer, organization of personal time, ability for adaptation to the social norms that allows one to adjust to his or her social environment. In graphology, layout and order help in determining how the writer stands before life in general [3, 7].

Dimension is the size and width of letters in a piece of writing. It is an objective parameter since it can be measured. It represents self-esteem of the writer, self-concept, how much an individual appreciates oneself, degree of introversion or extroversion, and degree of self-confidence [6, 7].

Connection shape is the structure per se of each and one of the letters in the writing (round, angled). The shape that letters take in a writing reflect appearance, modals, the way an individual conducts oneself,

whether one does it softly or intensively, with reservation or expansion, the culture and lifestyle. It provides an idea of the cultural, emotional, and mental level of a person. Figure 1 illustrates various connection shapes of handwriting [7].
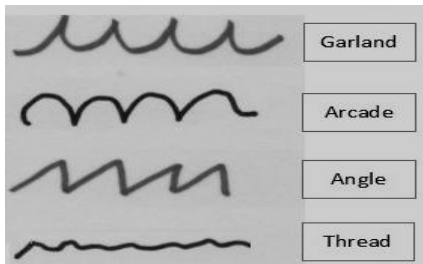

Fig 1: Different connection shapes of handwriting [7]

Slant or inclination indicates where letters are deviated to from the vertical edge. In order to study inclination, upper extensions and lower extensions are enlarged and the angle that is formed along with the vertical line is observed. From a psychological point of view, it expresses the need for an affectionate bond, need for contact, and dependency of an individual on others [7, 8].

Baseline indicates the direction of written lines. Examining a written page can help determine whether the baselines are oriented upwards or downwards, whether they stay horizontal or in serpentines on top of the line. Lines that are incomplete or include indentation are not considered. Traditionally, ascending lines are associated with optimistic people, whereas descending lines are associated with pessimistic attitudes [7, 9].

Speed of writing indicates the ability of comprehension, intellectual assimilation and the speed of mental processes. It points out the speed to resolve tasks. It allows evaluating the intellectual level, adjustment to new scenarios, temper, and activity. It also helps in assessing one's professional performance [5, 7].

Form or flow continuity represents the manner in which the letters tie between each other in one word. They can be connected, disconnected, grouped, or fragmented. Form thus indicates cohesion within the text [7].

Pressure is indicated by the strength of the mark printed on a sheet of paper. It symbolizes the intensity of the psychological energy of the writer. It represents the emotional security of an individual in the actions he or she performs and the degree of conviction one has in themselves. Figure 2 depicts samples of different writing pressures [7].
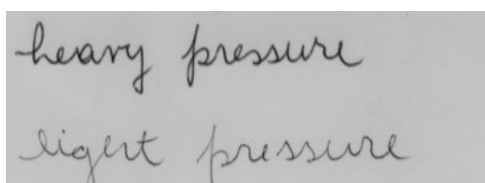

Fig 2: Different handwriting pressures [7]

Apart from handwriting recognition and analysis, text processing and mining aid in searching for repetitive patterns of writing that can be helpful in establishing an individual's characteristic style. While several software exist for processing text of digital documents, *optical character recognition* (OCR) software can recognize and convert handwritten documents into their digital counterparts for easier processing.

This paper thus proposes and discusses an algorithmic approach for analyzing human behavioral and psychological characteristics through text processing and handwriting recognition so that the obtained results can be incorporated with advanced artificial intelligence systems that can employ text processing and handwriting recognition as individualistic signatory features. Section II describes the proposed algorithmic approach in detail, along with the obtained results. Section III is devoted to result analysis. Finally, section IV elucidates the conclusions derived from the proposed study and future scope of work.

II. ALGORITHMIC APPROACH TO HUMAN BEHAVIORAL ANALYSIS

This section describes an algorithmic approach for analyzing human psychology through handwriting recognition and text pattern processing. In order to achieve this objective, three primary behavioral categories were defined with their associated handwriting attributes, as shown in Table I [7, 8, 9].

Table I
Behavioral categories with associated handwriting attributes

Behavioral Category	Handwriting Attributes
Category I: anger, depression, egoism, aggression, cruelty	Downward baselines, heavy pressure, large capital letters, leftwards slanting
Category II: generosity, calmness, amiability, adaptability, joy	Even pressure, wide spacing of words, rightwards slant, rising baselines
Category III: ambition, curiosity, concentration, individualism	Upright vertical script, small size, large capital letters, disconnected script

As described in Table I, Category I represents general negative personality traits such as anger and aggression. Category II represents general positive personality traits such as calmness and joy. Category III indicates general neutral personality traits such as curiosity and individualism. Each category has its own associated handwriting attributes that can help in assessing an individual's behavior. For the purpose of demonstration, Figure 3 illustrates two handwriting samples that were examined in the proposed study. Both handwriting samples belong to the same individual but clearly demonstrate how one's writing

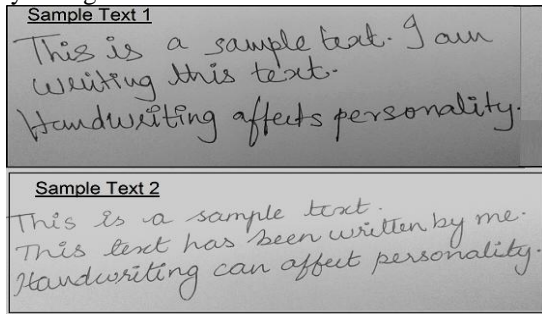style and patterns can change depending on the psychological state.

Fig 3: Handwriting samples

The two handwriting samples shown in Figure 3 were examined with respect to four parameters: pressure, slant, baseline, and dimension. MATLAB software image processing package was employed for carrying out the required handwriting analysis. TopOCR software was then employed for handwriting recognition and converting the images into their corresponding digital text, with subsequent text pattern processing with the help of RapidMiner software.

In order to determine the writing pressure, each text sample image was converted to grayscale, with pixel intensity values ranging from 0 (black) to 255 (white). Lower pixel intensity values of the writing indicates darker color, and thus heavy writing pressure. On the other hand, higher pixel intensity values indicate lighter color, and thus represent moderate to light pressure. The slant inclination of written characters and baseline directions were determined by dividing each text sample image into a grid of horizontal and vertical lines. Angular deviations of written characters from the vertical lines helped in determining the slant, whereas angular deviations of each written sentence from the horizontal lines helped in assessing the baseline directions. The character dimensions were measured and compared by examining how many cells in the text image grid were occupied by each letter. Through text processing with TopOCR and RapidMiner software, the overall tone of both the writing samples was assessed. Table II summarizes the obtained results for comparison.

Table II Writing attributes obtained from handwriting recognition and text processing

	Text Sample 1	Text Sample 2
Pressure	Heavy	Light and even
Slant	Leftwards	Rightwards
Baseline	Descending	Ascending
Dimension	Large capital letters	Even sized, regular spacing width
Writing Tone	Active voice	Passive voice

III. RESULTS ANALYSIS

The handwriting samples compared in the previous section reveal distinct personality traits and writing attributes, although both writing samples belong to the same individual. The first text sample was written by the participant individual while being in an aggravated and irritated psychological state, whereas the second text sample was written by the same individual while being in an elated and relaxed psychological state. Comparing Table I and Table II, it can be observed that the first handwriting sample demonstrates behavioral traits of Category I, whereas the second handwriting sample demonstrates behavioral traits of Category II. Furthermore, the usage of active voice in the first text sample and the usage of passive voice in the second text sample further highlight the psychological attributes of egoism (Category I) and amiability (Category II), respectively.

Although in the previous section two handwriting samples have been analyzed for the purpose of demonstration, the proposed algorithmic approach was applied on multiple other text samples, and in each case the obtained results have shown to agree with the known psychological state of the writer, thereby corroborating the proposed algorithmic approach.

IV. CONCLUSIONS

Human behavioral analysis has a multitude of applications, where tools such as handwriting recognition, graphology, and text processing can serve as crucial indicators of one's psychological state. This paper proposes an algorithmic approach for analyzing human behavioral and psychological characteristics through text processing and handwriting recognition. Three behavioral categories were defined with associated writing attributes. Two handwriting samples were examined with respect to the parameters of pressure, slant, baseline, dimension, and writing tone. The results obtained after processing each text sample agree with the known psychological state of the writer.

As future scope of work, greater number of categories can be defined for identifying subtle emotions and behavioral traits. Furthermore, apart from the four handwriting attributes incorporated in the proposed study, additional graphological features can be employed for an in-depth analysis of the written characters. The proposed algorithmic approach can also be extended for comparing text samples of multiple individuals, and text pattern processing can thereby be employed for identifying individualistic unique writing patterns.

REFERENCES

[1] Poorna Banerjee Dasgupta, Detection and Analysis of Human Emotions through Voice and Speech Pattern Processing, Vol. 52, No. 1, International Journal of Computer Trends and Technology, Oct. 2017.

[2] Javier G. Rázuri, David Sundgren, Rahim Rahmani, Antonio Moran Cardenas, Automatic Emotion Recognition through

Facial Expression Analysis in Merged Images Based on an Artificial Neural Network, IEEE 12th Mexican International Conference on Artificial Intelligence, Nov. 2013.

[3] Bradford R R, Bradford R B, Introduction to Handwriting Examination and Identification, Nelson-Hall: Chicago, IL, 1992.

[4] Erica Patino, Understanding Dysgraphia. [Online]. Available: https://www.understood.org/

[5] Klara G. Roman, Handwriting: A Key to Personality, First Edition, Pantheon Books, New York, 1952.

[6] B. Nevo, Scientific Aspects of Graphology: A Handbook, Springfield Publications, IL: Thomas, 1986.

[7] (2016) The Handwriting and Graphology website. [Online]. Available: http://www.handwriting-graphology.com/

[8] Beyerstein, Barry L, Beyerstein, Dale F, The Write Stuff: Evaluation of Graphology - The Study of Handwriting Analysis, First Edition, Buffalo: NY, Prometheus Books, 1992.

[9] Bayne, R., O'Neill, F., Handwriting and personality: A test of some expert graphologists' judgments, Guidance and Assessment Review, Volume 4, (p) 1–3, 1988.

AUTHOR'S PROFILE

Poorna Banerjee Dasgupta has received her B.Tech & M.Tech Degrees in Computer Science and Engineering from Nirma Institute of Technology, Ahmedabad, India. She did her M.Tech dissertation at Space Applications Center, ISRO, Ahmedabad, India and has also worked as Assistant Professor in Computer Engineering dept. at Gandhinagar Institute of Technology, Gandhinagar, India from 2013-2014 and has published several research papers in reputed international journals. Her research interests include image processing, high performance computing, parallel processing, and wireless sensor networks. Apart from research work, she is also currently engaged in editing and reviewing international journal research papers and articles for a prestigious multi-national firm.