Review on Recent Advances in Automatic Handwritten MODI Script Recognition

Kulkarni Sadanand A.
Vision and Intelligent System Lab
Department of CS & IT
Dr. BAMU University,
Aurangabad (MS) India

Borde Prashant L.
Vision and Intelligent System Lab
Department of CS & IT
Dr. BAMU University,
Aurangabad (MS) India

Manza Ramesh R.
Biomedical Image Processing Lab
Department of CS & IT
Dr. BAMU University,
Aurangabad (MS) India

Yannawar Pravin L.
Vision and Intelligent System Lab
Department of CS & IT
Dr. BAMU University,
Aurangabad (MS) India

ABSTRACT
Many HOCR systems was developed for different foreign language and efforts on Indian script still verdant. Due to the varying writing style of each individual HOCR becomes a critical challenge since last few decades. India was known to world for its unity in diversity. The geographical and cultural vastness of the country various spoken languages and writing scripts were invented and utilized for daily communication. Courageous endeavors was made towards HOCR for various Indian scripts. The development of HOCR system was considered an active area of research due to the complexity and structure of Indian scripts. MODI Script was an ancient script of India as compared with other Indian scripts. Large number of historical documents exists and they need to be preserved and explored therefore special attention must be given to MODI script. This research paper is an attempt towards an investigation of recent advances in Handwritten Optical Character Recognition system for MODI script.

Keywords
MODI Script, HOCR, Segmentation, Feature Extraction, Recognition, Image Classification

1. INTRODUCTION
Many researchers were attracted towards handwritten character recognition which is the most promising area of Pattern recognition, Image processing, Natural Language processing and Document analysis. HOCR was a process that provides intelligence to a computer that converts human writings into text. Due to advancements in technology pattern recognition and image processing lot of improvement was observed in recognizing handwritten characters.

Handwritten character recognition problem was defined based on general model of OCR with six stages 1) Preprocessing, 2) Segmentation, 3) Feature Extraction, 4) Classification and Recognition 6) Post processing. General framework of HOCR was shown in figure1.

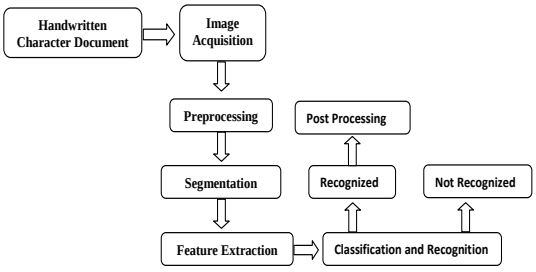

Figure 1: General framework of HOCR

Interest of this review was to analyze methodologies given in various six studies 1) Besekar D.N. [1] 2) Besekar D.N. [2] 3) Ramteke A.S. & Katkar G.S. [3] 4) Besekar D.N. & Ramteke R.J. [4] 5) Besekar D. N. & Ramteke R. J. [5] and 6) Ramteke A.S. and Katkar G.S. [6].

Rest of the paper was organized as, History of MODI script was given in section 2, Properties of MODI script was given in section 3, Pre-processing was discussed in section 4, Segmentation in section 5, Feature extraction methodologies in section 6, Techniques for classification and recognition in section 7, Post processing in section 8, Concluding the review in section 9, finally, Future research directions was discussed in Section 10.

2. HISTORY OF MODI SCRIPT
According to a people's linguistic survey of India conducted by David Lalmalsawma [7] India speaks 780 languages and 86 different scripts out of which 220 languages was disappeared in last 50 years and also bespeak a threat that 150 languages could disappear. MODI Script was one of the ancient script and vanished from official work during 19th century before Devanagari was officially adopted. Total time span where MODI was utilized for written communication at different stages was about 600 years. Meanwhile MODI became very popular and frequently used to write '*Marathi*' (primary language of Maharashtra in western India) [8,9].

Origin of this script still uncertain. According to Savant M.Y.[10], Rakesh A. Ramraje [11] MODI alphabets was developed by '*Hemadpant*' or '*Hemadri*' in 12th Century. *Hemadpant* was an administrator in the kingdom of '*Mahadev Yadav*' and '*Ramdev Yadav*' (1187-1318 at '*Devgiri*'). According to history expert Chandorkar, MODI script was evolved from *Mouryi (Bramhi)* script of *Ashoka* period. Besekar D.N. & Ramteke R.J. [4], Besekar D.N. & Ramteke R.J. [5] and Naren Ranadive [9] states that MODI alphabet was invented during the 17th century and frequently used for writing purpose in the era of *'Peshwai' (Pune)* and '*Chatrapati Shivaji Maharaj*'.

According to Salgaonkar *et.al.* [12] oldest available MODI document was of 1429 and 1389 A.D. preserved in *Bharat Itihas Sanshodhan Mandal (BISM) Pune*.

According to Anshuman Pandey [13] earliest in 12th century MODI script was known as '*proto-MODI*' or '*Adyakalin'*, 13th century '*Yadavakalin'*, '*Bahamanikalin*' of the 14th–16th century, '*Shivakalin*' of the 17th century and '*Chitnisi*' during 18th century called as '*Peshvekalin'* between 1818 to 1952.

3. PROPERTIES OF MODI SCRIPT

The MODI was written by '*Boru*' *or* '*lekhan*'. *'Lekhan'* was a pen created with the help of '*Bambuu*' which need to lift too often for dipping in the ink. The MODI script was formed by 46 distinctive letters in which 36 were consonants and 10 vowels. These characters were known as basic characters.

As compare to Devanagari there were 48 distinctive letters includes 36 consonants and 12 vowels. Long '*i*' and Long '*u*' was not used in MODI Script .

Before the commencement of writing in MODI characters a horizontal line was drawn across the page. Characters were written with respect to the horizontal line with full of moulds and curves to avoid lifting of *'Boru'* too often. According to Savant M.Y. [10] no punctuation marks and conjuncts were used in this script. Termination symbol for sentence or word was not used in MODI Script. Speed of writing was increased due to elimination of these symbols.

Consonants used in MODI script was shown in figure 2 and Vowels used in MODI Script was shown in figure 3.

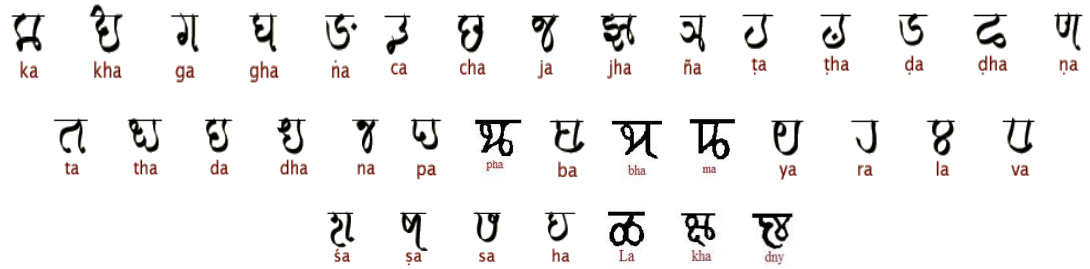

Figure 2: The consonants used in MODI Script.

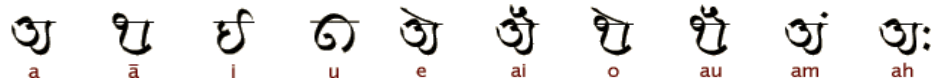

Figure 3: Vowels used in MODI Script.

Vowels was written as independent letters, or also by using diacritical marks at above, below, before or after to the consonant to which they belongs. Pattern of writing the vowels was known as modifiers and newly formed characters were called as conjuncts. Use of these modifiers was shown in figure 4 and some conjunctions were shown in figure 5.

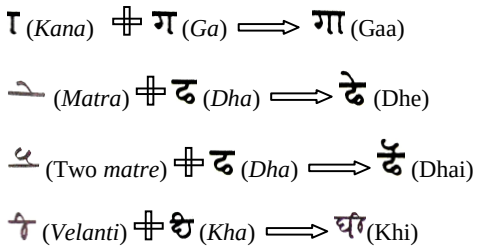

Figure 4: Use of Modifiers in Historical MODI Script

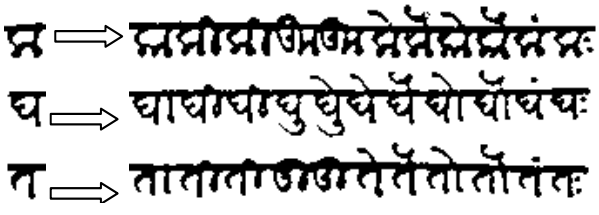

Figure 5: Conjunctions in Historical MODI Script

Sometimes two or more consonants were combined to produce new shape called compound character. Similar structural representation of characters was found in MODI, Devanagari, Hindi, Sanskrit and Marathi. Some examples of compound characters was shown in figure 6.

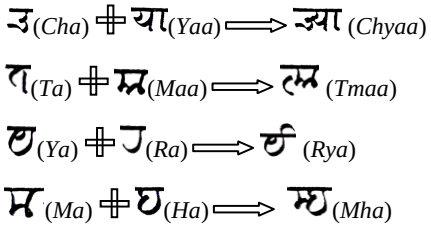

Figure 6: Compound characters in MODI Script

Pattern recognition of MODI Character was very challenging due to the inconsistency in the writing style, similarity in the character shapes, presence of modifiers and various other features of MODI scripts. The actual differences were visible in the behaviors of characters in certain environments, such as calligraphic uniformity of the characters, cursive nature of the script and degraded document. Similarity in calligraphic representation of 38 basic characters was shown in figure 7.

Figure 7: Uniformity in calligraphy of MODI characters

Principal interest of this review was to analyze different methodologies implemented in HOCR with special attention to MODI script. Despite of advancement in technology, HOCR for MODI Script was still in initial phase. Up till now only six studies were available for MODI script. Besekar D.N.[1] stated a morphological approach for recognition of five numerals from MODI script. Feature set was created by using blobs, vertical lines, horizontal lines, concavities present in the numerals. Recognition and classification was done by using decision tree and mathematical morphology. In this case 75% recognition rate was achieved for 0 and 1 and lower for 4,7 and 9.

Chain code and image centroid based recognition model was presented by Besekar D.N [2] for vowels in MODI script. In this work median filter was used for removing noise, global threshold for binarization, flood fill for avoiding boundary breaks and size normalization. Two layer feed forward Neural Network and SVM was used for classification and achieved 65.3% to 73.5% recognition rate.

Structural similarities was used by Ramteke A.S. & Katkar G.S [3,6] for recognition of MODI characters. Structural similarity approach was used for measuring image quality and image quality metrics. measured structure similarity (SSIM), KNN and back propagation NN was used for classification and achieved 91 to 97% recognition rate.

Zone based approach was used by Besekar D.N. & Ramteke R.J. [4] for recognition of offline handwritten numerals. In this study preprocessing was done by using median filter for removing noise, global threshold for binarization, flood fill for avoiding boundary breaks and size normalization. Feature set was created with the help of four equal square zones of size 15×15 and their polar coordinate, Variance, Theta angle and Rh distance. Using variance table for classification this study achieved 93.5% recognition rate.

Theoretical analysis of MODI Script recognition was done by Besekar D.N. & Ramteke R.J. [5]. Devanagari, MODI and Roman scripts were compared in this work and found that structural features were difficult to extract for MODI script. In this work internal and external segmentations were discussed and advised Internal segmentation for MODI script. In this analysis structural as well as topological features were suggested. This study also explains that HOCR for MODI script was a difficult task as compared to other handwritten script due to the cursive nature, variations in character, handwriting habits and synonymous structure of characters.

Summarily this review concludes that specific steps and results obtained in preprocessing was not shown by [1,3,4,5,6]. Methods and results of segmentation was unclear in [1-6] and neither of these studies were considered post processing for MODI script. Analysis of these six studies were shown in table 1**.**

Table 1. Comparison of results achieved in HOCR for MODI Script

Author and Reference	**Type of Data**	**Preprocessing**	**Segmentation**	**Feature Extraction**	**Classification and Recognition**	**Post Process**	**RR**
Besekar D.N. [1]	Five MODI numerals (0,1,4,7,9)	Morphological operations	Unclear	blobs, vertical & horizontal lines, concavities.	Mathematical morphology and Decision tree.	Unclear	75%
Besekar D.N. [2]	MODI vowels	Median filter Skewing, Binarization, Edge detection, Flood fill and Normalization	Unclear	Chain codes, Centroid, Normalized chain code histogram.	Two layer feed forward NN and SVM	Error correction and Unicode	65.3% to 73.5%
Ramteke A.S. & Katkar G.S. [3]	MODI characters	Depends on method of data acquisition	Unclear	Structural similarity approach	SSIM and KNN, back propagation NN.	Unclear	91 to 97%
Besekar D.N. & Ramteke R.J. [4]	MODI numerals	Morphological operations, Median filter Skewing, Binarization, Edge detection, Flood fill and Normalization	Unclear	Four square zones of size 15×15. Polar coordinate of zone, Variance, theta angle and Rh distance.	Variance table	Unclear	93.5%
Besekar D.N. & Ramteke R.J. [5]	MODI characters	Unclear	External and Internal segmentation	Structural and Topological features	Unclear	Unclear	--
Ramteke A.S. & Katkar G.S [6]	MODI characters	Depends on method of data acquisition	Unclear	Structural similarity approach	SSIM and KNN, back propagation NN	Unclear	91 to 97%

4. PRE-PROCESSING

Preprocessing was used to produce clean and inflectionless document to be used in later stages of HOCR. Various infections were noticed in old MODI historical documents. Plenty of such historical documents were about 600 years old [8,10] preserved with the help of cloth covers [1-6]. Keeping all these factors preprocessing stage became essential for the development of MODI HOCR. An review on offline handwritten character recognition was done by Nafiz Arica *et.al.* [14] and stated various essential preprocessing steps like noise removal, filtering, morphological operations, noise Modeling, normalization, Skew normalization and baseline extraction, slant normalization, size normalization, contour smoothing, compression, thresholding and thinning. Generally following techniques were used in various studies for preprocessing.

4.1 Noise Reduction

Noise removal was a process to remove noises produced at the time of capturing or transmission of the image. In case of historical MODI documents many types of noises were

introduced by scanning process such as disconnected line segments, bumps and gaps in lines, filled loops, local variations, rounding of corners, dilation, erosion, spike noise, random noise or independent noise, salt and pepper noise, black and white dot and dust particles. J. Serra [15] stated that it was necessary to eliminate various imperfections for best possible results in OCR.

4.2 Filtering

Filtering was a process to remove unwanted component from image. These unwanted components were produced by uneven writing surface, poor sampling rate of the data acquisition device. The book written by Ramesh Jain *et.al.* [16] stated various applications of image filtering techniques. Rohit Verma & Jahid Ali [17] stated use of filtering techniques and found that adaptive filter BM3D was the best choice for removing salt and paper noise also shows that median filter was reliable choice for filtering. Filtering techniques like smoothing, sharpening, threshold, contrast adjustment and textured or colored background elimination were studied by [18-22].

4.3 Morphological Operations:

Morphological operations are playing very important role for enhancing the accuracy of HOCR. Erosion, Dilation, Closing, Opening and skeleton, thinning were basic morphological operations. Although unexpected hideousness was produced by careless use of Morphological techniques the good quality of image was produced by careful use [14].

4.4 Skew angle detection

Inconsistency of scanning process and writing style were major constrains in HOCR for various languages. In MODI script writing style of characters were slightly tilted, curved and overlapped within the image and calligraphy of MODI script hurts accuracy of HOCR [6]. Therefore skew angle detection and correction became essential for HOCR in MODI script. Naazia Makkar & Sukhjit Singh [23] stated that projection profile, Hough transformation, K-NN Clustering, Thinning and Cross Co-relation for skew angle detection and correction.

4.5 Normalization

Normalization process used to modify all characters in a unique and standard form. Slant normalization and Size Normalization were commonly used methods [14]. Slant normalization focuses on angle of characters and Size Normalization focuses on size of the character. Due to calligraphic effect normalization become essential for MODI HOCR.

It was assumed that the preprocessing stage gives clear and normalized document with sufficient amount of shape information and high compression.

5. SEGMENTATION

Segmentation, divides the document into subcomponents like words, lines and characters. Accuracy produced by segmentation directly affect to the recognition rate. Segmentation became major concern area in many HOCR. Still there was a room for research in segmentation of touching, overlapping, cursive and complex scripts.

T. Steinherz *et.al.* [24] stated various segmentation free and segmentation based methods for cursive script. Abhishek Phukan & Mrinaljit Borah [25] stated two types of segmentation.

5.1 External Segmentation

External segmentation was used to extract paragraphs, sentences, or words with the help of structural analysis and functional analysis. Structural analysis was used to divide the document into blocks and functional analysis was used to find location and size of the components.

5.2 Internal Segmentation

Internal segmentation was used to extract isolates letters from cursive script. Internal segmentation was divided in Explicit Segmentation, Implicit Segmentation and Hybrid Segmentation [26,27,28]. Explicit Segmentation was used to find interconnections and cut the image according to the interconnections. In case of implicit segmentation a class of model images was designed and compared with the image. Image was then segmented according to match found in system search.

6. FEATURE EXTRACTION

Feature extraction process plays very crucial role in development of promising HOCR systems. Large no of feature extraction methods have been reported by researchers working on HOCR and effectiveness of these methods has been assessed in different studies using different forms of data. The book written by Mark Nixon and Alberto Aguado [29] stated that feature was a point of interest according to image classification and recognition also stated various feature extraction techniques. A review of feature extraction was done by Dong ping Tian [30] which summarizes features according color histogram, color moments, texture, shape, moments, scale and shape transformation.

Oivind Due Trier *et.al.* [31] stated use of template matching, deformable templates, unitary image transforms, zoning, geometric moment invariants, zernike moments and orthogonal moments for feature extraction. Gurpreet Singh *et.al.* [32] evaluated use of distance profile features, projection histogram features, zoning density features, zernike moment features and hybrid features for Gurmukhi Script.

Pritpal Singh & Sumit Budhiraja [33] stated zoning density with background directional distribution features for recognition of Gurumukhi script.

7. CLASSIFICATION AND RECOGNITION

Various techniques were developed since last few decades for classifying an unknown sample into a predefined class. Generally classification and recognition in HOCR was done by using holistic or analytic strategies [34,35,36]. Top down approach was used in holistic strategy for recognition of full word and segmentation problem was eliminated in this case. On other hand bottom up approach was used in analytic strategy for recognition of characters. Character segmentation was major concerned in this case. Errors at segmentation level was reflected in classification and recognition.

Priyanka Sharma & Manavjeet Kaur [37] discussed statistical model, syntactical or structural model, template matching model and neural network based model for pattern recognition. and also discussed use of Linear discriminant analysis (LDA), Quadratic Discriminant Analysis (QDA), Maximum entropy classifier, Decision trees, Kernel Estimation & K-nearest neighbor (KNN), Naive Bayes classifier, Artificial Neural Networks (ANN) and Support Vector Machine (SVM) for pattern classification.

Charles C *et.al.* [38] evaluated various methods for online as well as offline handwritten character recognition with special interest to Chinese script also discussed various types of recognition methods for cursive script.

Various recognition and classification techniques were reviewed by Nafiz Arica *et.al.* [14] and Gaikwad V.A. & D.S. Bormane [39] like template matching, statistical techniques, structural techniques, Neural Networks (NNs) for offline handwritten character recognition. Ramzi Haraty & Catherine Ghaddar [40] stated Artificial Neural Network (ANN) for Arabic text recognition and classification and found that promising results were produced by two generalized feed forward neural network.

8. POST-PROCESSING

A post processor was an integral part of any OCR system. Errors produced in recognition was major concern for implementation of post processing stage. An important role of correcting errors and resolving ambiguities in OCR was given to post processing in HOCR. Unfortunately, HOCR was still imperfect as it occasionally produces syntax and semantic errors in output text.

Post processing was done at word level, at sentence level and at semantics level. The most common post-processing technique at the word level was dictionary look-up method where an dictionary was developed to compare generated output of the OCR [41]. A dictionary can produce spell check and provide some alternatives for the outputs produced by recognizer. Ultimately utilization of a dictionary for correcting the minor mistakes was became a solution for OCR systems [41,42,43].

Accuracy of the HOCR was increased when post processing was implemented by using semantic information instead of using this context information. whereas contextual information was very important for development of vocabulary and extracting the meaning of the isolated word or phrase or statement. Therefore the incorporation of context and shape information in all the stages of OCR systems was necessary for meaningful improvements in recognition rates.

This review concludes that many promising methodologies were developed for OCR and HOCR in last few decades for Indian scripts like English (cursive) [26], Arabic [27], Gurumukhi [32], Persian[35], Devnagari [36], Tamil [44], Telugu [45], Bengali [46], and Kannada [47], Urdu [48], Gujarati [49] and Malayalam[50,51].

System architectures of OCR techniques for these studies were stated in Table 2. Although this work was done for different languages and can be considered for further enhancing of HOCR in MODI script.

Table 2. Existing OCR system architectures for various languages.

Author and Reference	Language	Preprocessing	Segmentation	Feature Extraction	Classification and Recognition	Post Process	RR
Amjad rehman *et.al.* [26]	English (cursive)	Digitization, Noise filter and Thinning	Implicit and Explicit segmentation	Hybrid features with Statistical features and Profile projection features	ANN	Lexicon	79.23% to 80.91 for Segmentation
Laslo Dinges *et.al.* [27]	Arabic	Median filter, Baseline estimation, Binarization, Thinning	Holistic approach	Structural and statistical features	Decision tree and Probability function	Lexicon	82% to 99.22% for Arabic words
Gurpreet Singh *et.al.* [32]	Gurumukhi	Morphological operations, Skew detection and correction, Noise removal	Zone based word segmentation	Distance Profile, Projection Histogram, Zoning density, Zernike moments and Hybrid feature	SVM, KNN	--	72.83% to 99%
Mir Mohammad Alipour [35]	Persian	8 bit-Gray Scale image, Binarization, Gaussian filter, Spatial filter for Smoothing and Thinning	Horizontal and Vertical projection profile	Structural features, Statistical features and Global Transformations.	Decision tree and Holistic (Global) or Analytic strategies	--	72 % to 98% Average RR 90.56%
Gunvantsinh Gohil *et.al.* [36]	Devnagari	Binarization, Noise removal, Scaling, Skew detection & correction,	Projection profile and graph based, Hierarchical approach	Binary features, Freeman Chain Code, PCA and FDA	ANN, SVM	--	54.10 to76.90% for ANN 60.02 to84.41% for SVM
C.Sureshkumar and Dr.T. Ravichandran [44]	Tamil	Noise removal, Skew detection and correction, smoothing, foreground extraction by threshold, Skeletonization, Thinning	Horizontal and Vertical projection profile	Height, width of character, no of horizontal and vertical lines, no of circles, no of arcs, centroid, pixels and Position in region.	SVM, Self Organizing maps (SOM), Radial Basis Function (RBF), RACHSU and Fuzzy NN	--	100% for train set 97% for test set.
Ramya J. and B. Parvathavarthini [45]	Telugu	Skew detection and correction, Binarization	Horizontal and Vertical projection profile	Structural features	Feed Forward Back Propagation Neural Network	--	90%
Farjana Yeasmin Omee [46]	Bengali	Binarization, Noise removal, Skew detection & correction, Page Layout	Horizontal and Vertical projection profile	Structural features	Decision tree, MLP, Kohonen NN, HMM	Spell checking & editing	98%
H.Imran Khan *et.al.* [47]	Kannada	Binarization, Thinning Noise removal	--	Chain code	*k*-Nearest Neighbor	--	100%
Tabassam Nawaz *et.al.* [48]	Urdu	Binarization, Noise removal, Filtering	Projection profile	chain code, shape and visual properties	Template matching	Unicode	89%
Brijesh Sojitra &Vishnukumar Dhakad [49]	Gujarati	Binarization, SOM	--	Clusters and mean value	*k*-Nearest Neighbor	--	97.78% for consonants
Rosemol Emmanuel and Jilu George [50]	Malayalam	Binarization, median filtering, Skew correction, Sobel edge detection	Horizontal and Vertical projection profile	Template	Template matching	Linguistic rules	--
Jomy John *et.al.* [51]	Malayalam	Noise Removal, Binarization, Edge detection, Normalization	Morphological operations	4 and 8 connected Chain code	A two layer feed forward neural network & SVM	error correction & mapping Unicode	61.2% to 72.1%

9. CONCLUSION

Principal objective of this research work was to analyze different methodologies used in HOCR with special attention to MODI Script. Regardless of paucity of the script and availability of the documents various problems like color representation, degradations, noise, calligraphy and boldness of characters were produced and can be tackled in preprocessing stage.

Segmentation was most studied and most concern part in various HOCR. Critical problem was faced for Segmentation in case of MODI script due to its cursive nature, lot of variations in writing style, overlapping lines, overlapping characters, presence of erratic '*shirorekha*'. Even though, technology was developed, overlapping character and overlapping line segmentation remained mysterious for HOCR and OCR.

Plenty of feature extraction methodologies were invented and invention was continued in this sense. Feature extraction for MODI HOCR was in infancy stage as compare to the amount of work done in various Indian or non Indian scripts. This review concludes that only structural similarity approach and one pattern of zoning was insufficient for reliable and accurate implementation of HOCR in MODI Script as there was various Low level, Shape and Size based, Region based, Texture based, Moment based and Zone based feature extraction methods were available.

Lot of work was done for classification using ANN, SVM, KNN, Bayesian classifier, and Decision Tree with respect to Indian and non Indian script. Although SSIM, SVM, KNN and NN were discussed for classification of MODI characters, implementation strategy and results produced for different environment was still unclear. Distinct representations of same character and contiguous representation of distinct character caused serious problems in classification of handwritten MODI script.

According to Rajesh Khillari [52], Wikipedia [53], Sattkaryotejak Sabha [54], The Times of India [55], The Hindu [56,57] and MID DAY INDIA NEWS [58] millions of MODI historical documents were preserved in Tanjavar's Saraswati Mahal, Oriental manuscript section of Chennai's Connemara University, Bharat Itihas Sanshodhan Mandal, Pune (BISM), Rajwade Sanshodhan Mandal, Dhule, museums in London, Paris, Spain, and Holland, South Asia, Europe, Denmark and other countries. These manuscripts were full of invaluable and priceless treasures from various fields like Literature, Science, Fine Arts, Ayurved, Pharmacy, Chemistry, Social Sciences, Psychology, Drawings, Paintings, Music, Astrology, History, Charms and Spells.

More scientific efforts were required to explore historical information stored in such a priceless treasure house by the development of HOCR for compound, conjunctions, numerals, and basic character sets of MODI script and various methods were yet to be implemented at different stages of HOCR for MODI script like segmentation, feature extraction, classification, recognition, and post processing.

10. FUTURE WORK

Research and development in MODI script was necessary to unfold the mysteries written on golden pages of history. An initiative was taken by Government of India for the development of Indian language by starting a program TDIL [59]. Main objective of this work was to study and develop various aspects of Indian languages.

CDAC, a well-known institute was also involved in the development of fonts and translators in collaboration with IIT Kanpur [60]. ISCII (Indian Scripts Standard Code for Information Interchange) [61] was developed by Indian government in 1988 for various Indian languages. Presently lot of room was available to do the research in the area of handwritten MODI script or character recognition form handwritten character documents or from handwritten MODI historical documents.

An example of handwritten MODI historical documents was shown in Figure 8.

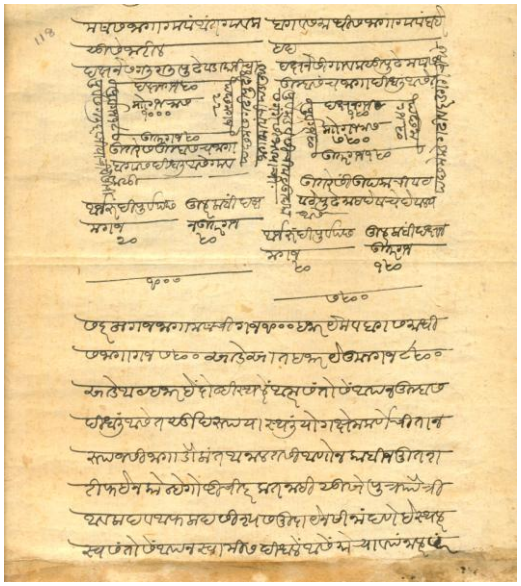

Figure 8: Handwritten MODI historical document.

11. REFERENCES

[1] Besekar D.N., 2011, "Recognition Of Numerals Of MODI Script Using Morphological Approach", International Referred Research Journal, ISSN 0974-2832, vol. 3, Issue -27, pp. 63-66.

[2] Besekar D.N., 2012, "Special Approach for Recognition of Handwritten MODI Script's Vowels", International Journal of Computer Applications (IJCA), MAHA-2012, pp. 48-52.

[3] Ramteke A.S., Katkar G.S., 2012, "Recognition of Offline MODI Script", International Journal of Research in Engineering, IT and Social Science (IJREISS), Volume 2, Issue 11, ISSN 2250-0588, pp.102-109.

[4] Besekar D.N., Ramteke R.J., 2012, "Feature Extraction Algorithm for Handwritten Numerals Recognition of MODI Script using Zoning-based Approach", International Journal of Systems, Algorithms & Applications, Volume 2, Issue ICRASE12, ISSN 2277 2677, pp. 1-4.

[5] Besekar D.N., Ramteke R.J, 2013, "Study for Theoretical Analysis of Handwritten MODI Script – A Recognition Perspective", International Journal of Computer Applications, vol. 64, no. 3, ISSN 0975-8887, pp. 45-49.

[6] Ramteke A.S., Katkar G.S. 2013, "Recognition of Off-line MODI Script- A structure Similarity Approach", International Journal ICT and management, Volume 1, Issue 1, ISSN 2026-6839, pp.12-15.

[7] David Lalmalsawma, 2013, "India speaks 780 languages, 220 lost in last 50 years–survey", available at "http://blogs.reuters.com/india/2013/09/07/india-speaks-780-languages-220-lost-in-last-50-years-survey", India Insights, Reuters, Edition US, Accessed: Mar 30,2015.

[8] Lawrence Lo, 2014, 'ancientscripts.com A compendium of world-wide writing systems from prehistory to today', "MODI", available at "www.ancientscripts.com/ modi.html", Accessed: Mar 30, 2015.

[9] Naren Ranadive, 2014, "The Origin and Development of Indian Writing System – MODI Script of Maharashtra", Randives Indian Antiquities and Inscriptions, available at "www.narendranath.webs.com", Accessed: Mar 30,2015.

[10] Savant M.Y., 2005, "Marathi MODI Script: Origin, Evolution and Significance", 'Kruti Rakshan', vol 1 no 1, Neha Paliwal ed(s), No. 5, Dr. Rajendra Prasad Road New Delhi 110 001, Mission Director National Mission for Manuscripts, page 16.

[11] Rakesh A. Ramraje, 2013, "History of MODI Script in Maharashtra", Global Online Electronic International Interdisciplinary Research Journal, Volume II, Issue 1, pp.54-58.

[12] Salgaonkar G.V., Chavan V.S., Mhatre K., Badkar D., Deshpande P., 2014, "History of MODI Lipi", available at "www.modilipi.com", Accessed: Mar 30,2015.

[13] Anshuman Pandey, 2011, "Proposal to Encode the MODI Script in ISO/IEC 10646", University of Michigan, Ann Arbor, Michigan, U.S.A.

[14] Nafiz Arica, Fatos T. Yarman-Vural, 2000, "An Overview of Character Recognition Focused On Off-line Handwriting", C99-06-C-203, IEEE.

[15] J. Serra, 1994, "Morphological Filtering: An Overview", Signal Processing, vol. 38, no.1, pp.3-11.

[16] Ramesh Jain, Rangachar Kasturi, Brian G. Schunch, 1995, "Machine Vision", McGraw-Hill, Inc. USA.

[17] Rohit Verma, Jahid Ali, 2013, "A Comparative Study of Various Types of Image Noise and Efficient Noise Removal Techniques", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, Issue 10, pp.617-622.

[18] Aditya Goyal, Akhilesh Bijalwan, Mr. Kuntal Chowdhury, 2012, "A Comprehensive Review of Image Smoothing Techniques", International Journal of Advanced Research in Computer Engineering & Technology, Volume 1, Issue 4, ISSN: 2278 1323, pp.315-319.

[19] Harvinder Singh, J.S. Sodhi, 2013, "Image Enhancement Using Sharpen Filters", International Journal of Latest Trends in Engineering and Technology (IJLTET), Vol. 2 Issue 2, ISSN: 2278-621X, pp. 84-94.

[20] Salem Saleh Al-amri, N.V. Kalyankar, Khamitkar S.D, 2010, "Image Segmentation by Using Threshold Techniques", Journal of Computing, Vol. 2, Issue 5, ISSN 2151-9617, pp. 83-86.

[21] Criminisi, P. P´erez, K. Toyama, 2004, "Region Filling and Object Removal by Exemplar-Based Image Inpainting", IEEE Transactions on Image Processing, Vol. 13, No. 9, pp.1-13.

[22] Bincy Antony M, K A Narayanankutty,2012, "Removing Occlusion in Images using Sparse Processing and Texture Synthesis", Int. Journal of Computer Science, Engineering and Applications,Vol.2,No.3, pp.117-124.

[23] Naazia Makkar, Sukhjit Singh, 2012, "A Brief tour to various Skew Detection and Correction Techniques", International Journal for Science and Emerging Technologies with Latest Trends, Vol 4, Iss.1, pp. 54-58.

[24] T. Steinherz, E. Rivlin, and N. Intrator, 1999, "Off-Line Cursive Script Word Recognition — A Survey," International Journal of Document Analysis and Recognition, vol. 2, no. 2, pp. 90–110.

[25] Abhishek Phukan, Mrinaljit Borah, 2014 "A Survey on Genetic Algorithm based Classification Technique for Handwritten Character Recognition", International Journal of Computer Applications (IJCA), 0975– 8887, National Conference cum Workshop on Bioinformatics and Computational Biology, NCWBCB-2014, pp 1-4.

[26] Amjad Rehman, Dzulkifli Mohamad and Ghazali Sulong, 2009, "Implicit Vs Explicit based Script Segmentation and Recognition: A Performance Comparison on Benchmark Database", International Journal of open Problems Compt. Math., Vol. 2, No. 3, September 2009 ISSN 1998-6262. pp 352-364.

[27] Laslo Dinges, Ayoub Al-Hamadi, Moftah Elzobi, Zaher Al Aghbari and Hassan Mustafa, 2011, "Offline Automatic Segmentation Based Recognition of Handwritten Arabic Words", International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 4, No. 4. pp 131-144.

[28] Munish Kumar, M. K. Jindal and R. K. Sharma, 2014, "Segmentation of Isolated and Touching Characters in Offline Handwritten Gurmukhi Script Recognition", International Journal of Information Technology and Computer Science, Vol 02, pp. 58-63.

[29] Mark Nixon, Alberto Aguado, 2008, "Feature Extraction and Image Processing", IInd Edition, Linacre house, Jorden Hill, London, UK.

[30] Dong ping Tian, 2013, "A Review on Image Feature Extraction and Representation Techniques", International Journal of Multimedia and Ubiquitous Engineering, Vollume 8, No. 4, pp.385-396.

[31] Oivind Due Trier, Anil K. Jain and Torfinn Taxt, 1996, "Feature Extraction Methods For Character Recognition A Survey",Pattern Recognition,Vol.29, No.4,pp.641-662.

[32] Gurpreet Singh, Chandan Jyoti Kumar, Rajneesh Rani and Dr. Renu Dhir, 2013, "Feature Extraction of Gurmukhi Script and Numerals: A Review of Offline Techniques", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, Iss.1, ISSN:2277 128X, pp.257-263.

[33] Pritpal Singh, Sumit Budhiraja, 2011 "Feature Extraction and Classification Techniques in O.C.R. Systems for Handwritten Gurmukhi Script – A Survey", International Journal of Engineering Research and Applications, ISSN: 2248-9622, Vol. 1, Issue 4, pp. 1736-1739.

[34] Vikas J Dongre and Vijay H Mankar, 2010, "A Review of Research on Devnagari Character Recognition", International Journal of Computer Applications, Volume 12– No.2, pp.8-15.

[35] Mir Mohammad Alipour, 2013, "A Novel Approach to Recognition of the Isolated Persian Characters using Decision Tree", International Journal of Computer Applications, Volume 66– No.12,, pp.14-20.

[36] Gunvantsinh Gohil, Rekha Teraiya and Mahesh Goyani, 2012, "Chain Code And Holistic Features Based OCR System For Printed Devanagari Script Using ANN And SVM", International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.1, pp.95-104.

[37] Priyanka Sharma, Manavjeet Kaur, 2013, "Classification in Pattern Recognition: A Review", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, Iss.4, 3 ISSN:2277 128X, pp.298-306.

[38] Charles C. Tappert, Ching Y. Suen, Toru Wakahara, 1990, "The State of The Art In On-Line Handwriting Recognition," IEEE Transciions On Pattern Analysis And Machine Lntelligence Vol. 12, No. 8, pp. 787–808.

[39] Gaikwad V.A., D.S. Bormane, 2008, "An Overview of Character Recognition Focused On Offline Handwriting", International Journal of Computer Science and Applications, Vol.1, No.3, ISSN 0974 1003, pp.157-159.

[40] Ramzi Haraty and Catherine Ghaddar, 2004, "Arabic Text Recognition", International Arab Journal of Information Technology, Vol. 1, No. 2, pp. 156-163

[41] Rajesh Khillari, 2002, "A Post-Processor for Gurmukhi OCR", Sadhana, Vol. 27, Part 1, pp. 99–111.

[42] Dharam Veer Sharma, Gurpreet Singh Lehal, Sarita Mehta, 2009, "Shape Encoded Post Processing of Gurmukhi OCR", 10th International Conference on Document Analysis and Recognition, IEEE, pp.788–792.

[43] Tao Hong and Jonathan J. Hull, 1995, "Visual Inter-Word Relations and their Use in OCR Post processing", Proceedings of the Third International Conference on Document Analysis and Recognition, IEEE, pp.442–445.

[44] C. Sureshkumar and Dr. T. Ravichandran, 2010, "Handwritten Tamil Character Recognition and Conversion using Neural Network", International Journal on Computer Science and Engineering, Vol. 02, No. 07 pp. 2261-2267.

[45] Ramya J. and B. Parvathavarthini, 2014, "Feed Forward Back Propagation Neural Network Based Character Recognition System For Tamil Palm Leaf Manuscripts", Journal of Computer Science, 10 (4), pp-660-670.

[46] Farjana Yeasmin Omee, Shiam Shabbir Himel and Md. Abu Naser Bikas, 2011, "A Complete Workflow for Development of Bangla OCR", International Journal of Computer Applications (IJCA) 0975–8887, Volume 21, No. 9, pp. 1-6.

[47] H. Imran Khan, Smitha U. V, Suresh Kumar D. S, 2013, "Isolated Kannada Character Recognition using Chain Code Features", International Journal of Science and Research, Volume 2 Issue 8, pp 67-70.

[48] Tabassam Nawaz, Syed Ammar Hassan Shah Naqvi, Habib ur Rehman & Anoshia Faiz, "Optical Character Recognition System for Urdu (Naskh Font) Using Pattern Matching Technique", International Journal of Image Processing, (IJIP), Vol. 3, Issue 3, pp 92-104.

[49] Brijesh Sojitra and Vishnukumar Dhakad, 2012, "Neural Network In Character Recognition Of Gujarati Script", Journal Of Information, Knowledge And Research In Computer Engineering, Vol. 02, Issue 02, pp 269-27.

[50] Rosemol Emmanuel and Jilu George, 2013, "Automatic detection and recognition of Malayalam text from natural scene images", IOSR Journal of VLSI and Signal Processing, Volume 3, Issue 2, PP 55-61.

[51] Jomy John, Pramod K. V, Kannan Balakrishnan, 2011, " Offline Handwritten Malayalam Character Recognition Based on Chain Code Histogram", Proceedings of ICETECT, 978-1-4244-7926-9/11, PP 736-741.

[52] Rajesh Khillari, 2008, "History of MODI Script", "Shri Modi Vaibhav",available at"modiscript.blogspot.in/2008 /05/history-of-modi-script.html",Accessed:Mar 30, 2015.

[53] Wikipedia, 2015, "Saraswathi Mahal Library", available at "en.wikipedia.org/wiki/Saraswathi_Mahal_Library", Accessed: Mar 30, 2015.

[54] Sattkaryotejak Sabha, 2014, "Vagdevata Mandir", available at "www.dasbodha.org/index.php?option =com_content&view=article&id =104&Itemid =652", Accessed: Mar 30, 2015.

[55] The Times of India, 2014, "Band of researchers, enthusiasts strive to keep Modi script alive", available at "www.timesofindia.indiatimes.com/city/ pune/Band-of-researchers-enthusiasts-strive-to-keep-Modiscriptalive/ articleshow/30761335.cms", Accessed: Mar 30, 2015.

[56] The Hindu, 2007, Tanjawar, "House library panel visits Saraswathi Mahal", available at "thehindu.com/todays-paper/tp-national/tp-tamilnadu/house-library-panel-visits-saraswathimahal/article1924551.ece", Accessed: Mar 30, 2015.

[57] The Hindu, 2013, Thanjavur (TN.), "MODI documents of Maratta era to be digitised", available at, "thehindu.com/todays-paper/tp-national/modi-documents -of-maratta-era-to-be-digitised/article4738336.ece", Accessed: Mar 30, 2015.

[58] Mid Day India News, 2013, Pune, Priyanka Deshpande, "Rare Manuscripts to be Digitalised by archives dept", available at "http://archive.mid-day.com/news/2013/feb/ 130213-pune-rare-manuscripts-to-be-digitalisedby archives-dept.htm", Accessed: Mar 30, 2015.

[59] Technology Development for Indian Languages (TDIL), 2015, "TDIL", available at "http://tdil.mit.gov.in", Accessed: Mar 30, 2015.

[60] Centre for Development and Advanced Computing (CDAC), 2015, "Fonts within an Indian Perspective", available at "www.cdac.in/index.aspx?id=mlc_gist_font" Accessed: Mar 30, 2015.

[61] Wikipedia, 2015, "Indian Script Code for Information Interchange", available at "http://en.wikipedia.org/wiki/ Indian_Script_Code_for_Information_Interchange", Accessed: Mar 30, 2015.