

Multiple Characters Matching in Document Image Using Handwriting Recognition

Pinagadi Venkateswararao¹ and S. Murugavalli²

¹Sathyabama University, Kamaraj Nagar, Semmancheri, Chennai – 600119, Tamil Nadu, India; venkat.pinagadi@gmail.com

²Department of Computer Science and Engineering, Panimalar Engineering College, Poonamallee, Chennai – 600123, Tamil Nadu, India; murugavalli26@rediffmail.com

Abstract

Objectives: To propose suitable system to filter the particular keyword in the document image. **Methods/Statistical Analysis:** Input image is pre-processed by using Gabor filtering. The preprocessed image is segmented by using segmentation free method. Then the feature extraction is done by using block based method. The Connectionist Temporal Classification (CTC) token passing calculation restores an arrangement of words from the lexicon. Then the word matching is performed by using normalized cross correlation. **Findings:** The performance of the proposed and existing methods are analysed with the help of word similarity, recognition accuracy, Recognition rate, computational time. Template matching method is utilized to discover and perceive the format picture which is found in the given info picture. Standardized cross connection techniques give precise outcome for the archive pictures. **Application/Improvements:** There are many applications such as data entry from printed paper data records, whether passport documents, invoices, bank statements, computerized receipts, business cards, mail, printouts of static-data, or any suitable documentation.

Keywords: Block Based Feature Extraction And Optical Character Recognition, Matching, Normalized Cross Correlation, Template Matching

1. Introduction

Acknowledgment of hand written characters is a testing zone in design acknowledgment and picture preparation. In spite of the fact that it has been the subject of research for quite a while, penmanship acknowledgment is as yet a broadly unsolved issue. Under troublesome conditions, for example, substantial vocabularies, diverse written work styles or corrupted archives, catchphrase spotting arrangements have been proposed rather than a total translation to spot words in report pictures.

Penmanship acknowledgment has been a standout among many captivating and examination of regions in research in field of picture assembling and framework recognition in the ongoing years. When all said is done, penmanship acknowledgment is characterized into

two sorts as disconnected and on-line penmanship acknowledgment techniques. In the disconnected acknowledgment, the written work is generally recorded visually by a scanner and the completed arrangement is available as a photo. However, for the two dimensional headings of dynamic centers are addressed as a segment of time and the demand of strokes in the on-line structure made by the author are likewise accessible¹.

By hand character acknowledgment, Image processing and pattern recognition are assumed as written in critical part. Transformation of transcribed characters is basic for making a couple of basic records related to our history, for instance, unique duplicates into machine editable edge, so that, it has a tendency to be easily gotten to and show free work is proceeding in Optical Character Recognition that is the treatment of printed/PC delivered report, written by

*Author for correspondence

hand and physically made archive preparing i.e. manually written character acknowledgment.

Acknowledgment and extraction of content in archive pictures is the point of report picture investigation. Anyway data recovery is worried about substance based report perusing, ordering and looking from a gigantic database of record pictures. The content recovery from report pictures has gained critical ground and tending to related data handling issues, for example, subject grouping and data separates.

Data recovery from record pictures has turned into a developing and testing issue. Data recovery frame report pictures are created utilizing two methods².

The principal approach is called acknowledgment construct recovery which is situated in light of Optical Character Recognition (OCR). OCR is a system which is utilized to distinguish the characters from record pictures, at that point changes over these pictures into their content organization. After transformation, archives can be altered, looked and put away³.

The second approach sans acknowledgment recovery which depends on Word spotting procedure and it is utilized to coordinate and recover data from archive pictures with no transformation. It finds the client determined watchword from archive picture by a word-to-word coordinating⁴. Word spotting strategy played out its errand by utilizing division. This procedure includes the accompanying strides for seeking watchwords from archive picture. They are preprocessing, edge discovery, and division, include extraction and coordinating.

In^{5,6} proposed word spotting structure for getting to the substance of chronicled machine-printed records, without the utilization of an optical character acknowledgment. The proposed philosophy has assessed the authentic Modern Greek printed reports which was accessible amid the seventeenth and eighteenth century. Character-based word demonstrates makes it conceivable to operate ASCII questions, subject to the accessibility of an arrangement of named model characters⁶. To enhance the productivity of getting to and seeking, characteristic dialect handling systems have been tended to. Looking through the report pictures utilizing just a base word-frame for finding all the comparing bent word-shapes and an equivalent word lexicon that further encourages access to the semantic setting of archives⁶.

In⁷ determined a novel system for word spotting in transcribed record pictures. It utilizes neighborhood

closeness look without utilizing any preparation information. They utilized both division based word spotting and division free word spotting strategies. Four recorded written by hand informational indexes are utilized for explore investigation utilizing standard assessment measures. It is discovered that the proposed calculation gives preferable execution over existing word spotting methods.

In⁸ displayed an exhaustive report on word spotting method for different contents or text styles. They broke down a different word spotting procedures for coordinate the catchphrase. Depicted the means in particular preprocessing - highlight extraction, portrayal and likeness measures to recover data from the archive pictures. Creators additionally talked about generally utilized datasets for word detecting, the best in class in the most regularly utilized record picture databases. At last they close learning based word spotting procedure gives preferable outcomes over adapting free technique. Learning-based techniques without requirement for physically recover to looking through the content inquiries.

Word spotting strategies might be partitioned into numerous classifications as per different variables. Contingent upon how the information is indicated by the client we can recognize question by-case (QBE) from inquiry by-string (QBS) strategies. In the QBE situation, the client chooses a picture of the word to be looked in the record gathering, while in the QBS worldview, the client gives a self-assertive content string as contribution to the framework Alternate approach to arrange word spotting techniques relies upon in the case of preparing information are utilized disconnected, either to learn character and word models or tune the parameters of the framework. Along these lines we can recognize taking in based from adapting free methodologies. At last, word spotting techniques which can be specifically connected to entire record pages are considered as division free, interestingly with division based strategies, where a division step must be connected at line or word level amid preprocessing⁹.

Catchphrase spotting approaches are comprehensively delegated Character based watchword spotting and Word based spotting and in some cases there is a joining of the two¹⁰. On basis of spotting in character, highlights are spoken to at level of character. The character that is: encoded with shape code of every character and

can be added from word to speak. On basis of spotting in Word, highlights are removed at the level of word specifically rather than character level which is unfeeling to character division blunder. From this perspective, on basis of word spotting in character, the strategies can be delegated Character Shape Analysis techniques, on basis of character protest strategies in N-gram and on basis of pixel in strategies of character coordinating.

Creator displays a completely unsupervised surface division calculation by utilizing adjusted discrete wavelet outlines decay and a mean move calculation. By completely unsupervised, we mean the calculation does not require any information of the sort of surface present nor the quantity of surfaces in the picture to be divided.

The essential thought of the proposed technique is to utilize the changed discrete wavelet casings to separate valuable data from the picture. At that point, beginning from the least level, the mean move calculation is utilized together with the fluffy c-implies grouping to partition the information into a proper number of bunches. The information bunching process is then refined at each level by considering the information at that specific level. The last firm division is acquired at the root level. This approach is connected to fragment an assortment of composite surface pictures into homogeneous surface regions and great division results are accounted for¹¹. This prompts a normal right grouping up to 96%. The experimentation gets rehashed for various resistance levels both for Brodatz, and the Vistex pictures. The technique performed extremely well in tests. It isn't touchy to the choice of parameter esteems, does not require any earlier learning about the quantity of surfaces or locales in the picture, and appears to give essentially preferred outcomes over existing unsupervised surface division approaches.

2-D Gabor channel is a prominent instrument in restorative picture order^{12,13}, surface investigation and separation. The procedure of surface division utilizing Gabor filters¹² includes a legitimate channel bank outline that ought to be altered for various frequencies of spatial and introductions to hide up the spatial recurrence space, breaking down the picture into various sifted pictures; highlight extraction from these pictures, and grouping of the pixels in the element space to deliver portioned picture. On producing surface highlights utilizing multichannel channels two essential issues must be tended to. The principal issue manages the useful portrayal of the channels and in addition their number, introduction

and separating. The second issue manages separating noteworthy highlights by information combination from various channels. Surface division requires concurrent estimations in both the spatial and the spatial recurrence areas. Channels with littler transfer speeds in the spatial recurrence space are more attractive as they enable us to do fine qualifications.

2. Research Methodology

The proposed Methodology aims in filtering the particular keyword in the document image as shown in Figure 1.

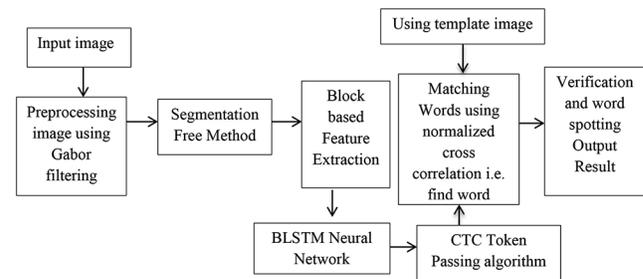


Figure 1. Proposed methodology.

2.1 Block Based Feature Extraction

Standardized record picture was part into a few word cases, so as to catch change varieties as far as turn and scaling. This square based element extraction computes distinctive arrangement of highlight vectors in light of ascertaining 5×5 non-covering pixel densities and applying word picture interpretation⁵.

3. Segmentation Free Method

In word spotting strategy without performing division is called as division free technique. Division free word spotting method has the accompanying strides for looking watchwords from record picture. They are Preprocessing, Normalization, Salient district location, square based component extraction, Candidate picture territories, Detection of word occasions and Remove Overlapping result.

4. Bi-directional Long Short Term Memory (BLSTM)

The neural system depicts the grouping probabilities of each letter and each situation in the following line. This

arrangement can be productively utilized in word and next line acknowledgment and in addition catchphrase acknowledgment. The CTC token passing calculation is in conjunction with BLSTM neural system.

5. Passing Algorithm for CTC Token

The Connectionist Temporal Classification (CTC) token passing calculation restores an arrangement of words from the lexicon whose score is (locally) ideal. CTC is best utilized in conjunction with design equipped for joining long-extend setting in both info headings. A calculation, in view of the token passing calculation that enables us to locate a surmised answer for a basic language. The CTC Token Passing calculation takes this grouping of letter probabilities, and additionally a lexicon and a dialect demonstrate, as its info and figure an imaginable succession of words.

This calculation works in yield layer. The CTC token passing calculation portrayed in takes as its information the yield initiations of the neural system and factual data about every unmistakable word, which infers that a lexicon manages which words, can be perceived by any stretch of the imagination. Because of the acknowledgment procedure, we get the translation of the given content line, i.e. a feasible arrangement of words.

The Token Passing of CTC calculation for single word expects an arrangement of letter probabilities of length t as yield by the system, together with the word was a succession of ASCII characters, and returns a coordinating score i.e. the likelihood that the contribution to the system was to be sure the given word. Let n (m , p) mean the likelihood of the letter l happening at the position k as indicated by the system yield and $a = l_1, l_2 \dots l_n$ signify the word to be coordinated. The calculation initially grows w into a succession.

$$w = \epsilon l_1 \epsilon l_2 \epsilon \dots \epsilon l_n \epsilon = c_1 c_2 \dots c_{2n+1}$$

what's more, makes each character c_p ($p = 1, \dots, 2m+1$) and each position $q = 1 \dots t$ in the following line of the token $\vartheta(p, q)$ to store the likelihood that character c_p is available at position q together with the likelihood of the optimal way starting to position q . The calculation steps are demonstrated as follows.

Require: input $w = l_1, l_2, \dots, l_n$

- 1: *Initialization:*
- 2: *expand* w to $w' = \epsilon l_1 \epsilon l_2 \epsilon \dots \epsilon l_n \epsilon = c_1 c_2 \dots c_n$
- 3: $\vartheta(1, 1) = n(\epsilon, 1)$
- 4: $\vartheta(2, 1) = n(l, 1)$
- 5: *Main loop:*
- 6: *for all sequence positions* $2 \leq j \leq t$ *do*
- 7: *for all positions* i *of the extended word* $1 \leq i \leq 2n+1$ *do*
- 8: $T_{best} = \{\vartheta(i, j-1)\}$
- 9: *if* $i > 1$ *then*
- 10: $T_{best} = T_{best} \cup \vartheta(i-1, j-1)$
- 11: *if* $c_i \neq \epsilon$ *and* $c_i \neq c_{i-2}$ *then*
- 12: $T_{best} = T_{best} \cup \vartheta(i-2, j-1)$
- 13: *end if*
- 14: *end if*
- 15: $\vartheta(i, j) = \max(T_{best}).n(i, j)$
- 16: *end for*
- 17: *end for*
- 18: *return* $\max\{\vartheta(2n+1, t), \vartheta(2n, t)\}$

For each word w , the changed word w' as w with spaces included toward the start and end and between each match of names.

6. Template Matching using normalized Cross Correlation

Template matching technique strategy can be connected just with the assistance of layout pictures. It is utilized to discover the pixel level and matches the character limits in the layout picture. Numerous layout coordinating systems are utilized in picture preparing, which are utilized to discover the area of the hunt picture.

6.1 Cross Correlation

The cross-connection layout coordinating is persuaded by the separation measure (squared Euclidean separation)

$$d_{f,n}^2(s, q) = \sum_{a,b} [f(a, b) - n(a-s, b-q)]^2 \quad (1)$$

Where f is the info picture and t is the layout picture, the entirety is over a, b under the window containing the element n situated at (s, q) in the extension of.

$$d_{f,n}^2(s, q) = \sum_{a,b} \left[f^2(a, b) - 2f(a, b)n + n^2(a-s, b-q) \right] \quad (2)$$

The term $\Sigma t^2(a-s, b-q)$ is constant. If the term $\Sigma f^2(a, b)$ is around constant, then the persisting term of cross-correlation.

$$c(s, q) = \sum_{a,b} f(a, b) n(a-s, b-q) \tag{3}$$

Eq. (3) represents the identity measures among image and the feature.

6.2 Cross Correlation of Normalization

There are a few hindrances for utilizing the Eq. (3) for format coordinating

- If the picture vitality shifts with position $\Sigma f^2(a, b)$ coordinating utilizing Eq. (3) can fizzle. For instance, the connection amongst element and a precisely coordinating area in the picture might be not as much as the relationship between's the component and a brilliant spot.
- The scope of $c(s, q)$ is subject to the extent of the element.
- Eq. (1) is not invariant to changes in picture sufficiency, for example, those caused by altering lighting conditions over the picture grouping.

The relationship coefficient conquers these challenges by normalizing the picture and highlight vectors to unit length, yielding a cosine-like connection coefficient.

7. Performance Analysis

The execution investigation of proposed work is the different characters coordinate the section in which we need to check whether the given word is found or not. The parameters are Word likeness, Recognition precision, Recognition rate, Computational runtime. Existing arrangement of cross connection technique is contrasted and standardized cross relationship strategy.

7.1 Computational Runtime

Correct by and large runtime is hard to quantify because of the way that diverse projects are utilized, composed in various dialects, and not the whole horse streamlined for speed as shown in Figure 2.

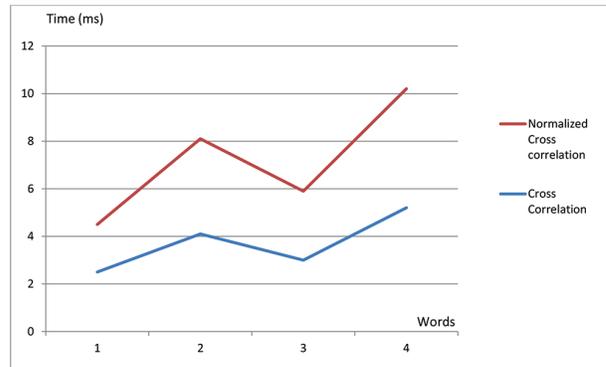


Figure 2. Computational time.

7.2 Word Similarity

The word similitude is computed by the word, they looked word which is discovered it out effectively from the gathering of words as shown in Figure 3.

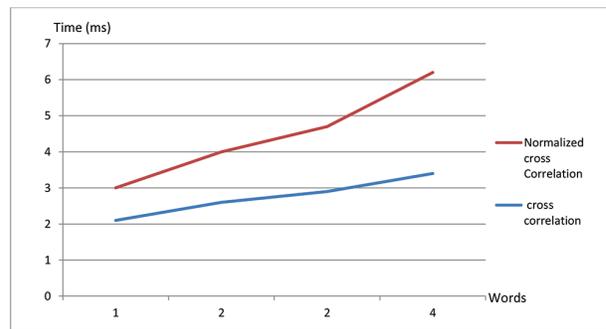


Figure 3. Similarity word.

7.3 Recognition Accuracy

Accuracy is ascertained in rate which is accurately perceived picture by the entire report picture in the framework as shown in Figure 4.

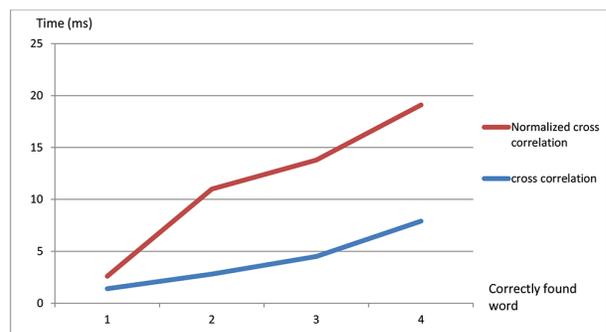


Figure 4. Recognition accuracy.

7.4 Recognition Rate

Rejection rate is computed by the word which isn't coordinates in client inquiry and the entire record picture in the framework as shown in Figure 5.

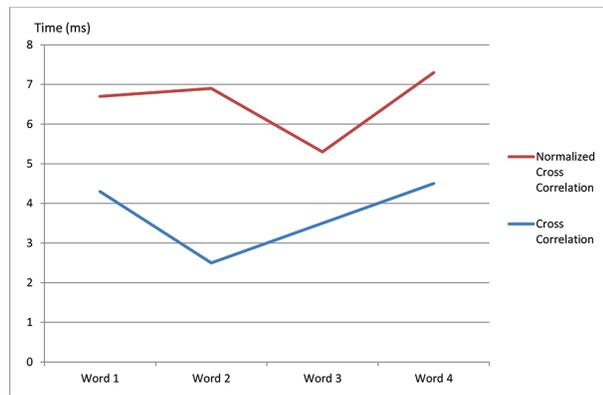


Figure 5. Recognition rate.

8. Conclusion

Template matching method is utilized to discover and perceive the format picture which is found in the given info picture. Standardized cross connection techniques give precise outcome for the archive pictures. Many testing research issues are accessible in checked archive pictures. These issues can be tackled by growing new calculations, ideas and strategies.

9. Reference

- Pradeep J, Srinivasan E, Himavathi S. Neural Network based Handwritten Character Recognition system without feature extraction. International Conference on Computer, Communication and Electrical Technology; 2011. p. 40–44. <https://doi.org/10.1109/ICCET.2011.5762513>.
- Vijayarani S, Sakila A. A survey on word spotting techniques for document image retrieval, International Journal of Engineering Applied Sciences and Technology. 2016; 2(2):6–10.
- Vijayarani S, Sakila A. Performance comparison of OCR tools, International Journal of UbiComp. 2015; 6(3):19–29. <https://doi.org/10.5121/iju.2015.6303>.
- Letta A. Feature extraction and Matching in Amharic Document Image Collections. 2011.
- Gatos B. Segmentation-free Word Spotting in Historical Printed Documents. 10th International Conference on Document Analysis and Recognition; 2009. <https://doi.org/10.1109/ICDAR.2009.236>.
- Kesidis AL, Galiotou E, Gatos B, Pratikakis I. A word spotting framework for historical machine printed documents. IJDAR, Springer-Verlag; 2010. p. 1–14. <https://doi.org/10.1007/s10032-010-0134-4>.
- Zagoris K, Pratikakis I, Gatos B. Unsupervised word spotting in historical handwritten document images using document-oriented local features, IEEE Transactions on Image Processing. 2017; 26(8):4032–41. <https://doi.org/10.1109/TIP.2017.2700721>. PMID: 28475054.
- Angelos P, Giotis, Giorgos Sfikas, Basilis Gatos, Christophoros Nikou. A survey of document image word spotting techniques, Pattern Recognition. 2017; 68:310–32. <https://doi.org/10.1016/j.patcog.2017.02.023>.
- Rodr JA, Serrano G, Perronnin F. Handwritten word-spotting using hidden Markov models and universal vocabularies, Pattern Recognition. 2009; 42(9):2106–16. <https://doi.org/10.1016/j.patcog.2009.02.005>.
- Lu S, Linlin L, Tan CL. Document image retrieval through word shape coding, IEEE Transactions on Pattern Analysis and Machine Intelligence. 2008; 30(11):1913–18. <https://doi.org/10.1109/TPAMI.2008.89>. PMID: 18787240.
- Hui Zhang, Fritts JE, Goldman SA. Image segmentation evaluation: A survey of unsupervised methods, Computer Vision and Image Understanding. 2008; 110:260–80. <https://doi.org/10.1016/j.cviu.2007.08.003>.
- Grigorescu SE, Petkov N, Kruizinga P. Comparison of texture features based on Gabor filters, IEEE Transactions on Image Processing. 2002; 11(10):1160–67. <https://doi.org/10.1109/TIP.2002.804262>. PMID: 18249688.
- Kop AM, Ravindra Hegadi. Kidney Segmentation from Ultrasound Images using Gradient Vector force. IJCA Special Issue on -Recent Trends in Image Processing and Pattern Recognition RTIPPR; 2010. p. 1–6.