Skewed Character Recognition using Permutation Invariant RAS Transform

A.V. Narasimha Rao
Professor, Department of ECE, Chaitanya Bharathi Institute of Technology, Hyderabad

M. Ramanjaneyulu
Associate Professor, Dept of ECE, KITE Professional College of Engineering Sciences, Shabad, Hyderabad.

M. Devaraju
Associate professor, Department of ECE, CMR Institute of Technology, Hyderabad

ABSTRACT

Several applications of cognitive informatics are coupled with automatic acquisition and processing of data and images for recognition and classification. There are many algorithms with wide-ranging performance in pattern recognition. A computer recognizes a character as a pattern from its digital image or from a set of parameters picked from its digital representation. This paper explores an optimal algorithm for pattern and printed character recognition which improves the performance in our world of noise and multiple styles of font. This approach is based on a permutation invariant Rao Alaka Shift Transformation (RAST) and tests for several printed alphanumeric characters in three different fonts. A template of the character for comparison is generated, in the transformed domain, after applying RAST to the digital image of the character in the chosen fonts. Pattern recognition is attempted also by rotating the images by about 25 degrees on either side. Better recognition performance is expected of the proposed approach when compared with the spatial domain. The results obtained in this study are found to be very encouraging and promising.

KEYWORDS

Permutation Invariant Transforms; Rao Alaka Shift Transform; Template matching; character recognition; Rotation invariance

1. INTRODUCTION

Pattern recognition techniques are often an important component of intelligent systems and are used for both decision making and data processing. A pattern can be understood as an arrangement of descriptors or a set of descriptors for a specific arrangement. In the most general sense, pattern recognition is a science that is related with detection of a specific arrangement or property in an object by making suitable measurements of descriptors and associating the recognized object with a group (class) of objects sharing several common properties.

The subjects like object recognition and classification belong to the broader field of machine intelligence- the methods of making machines learn and reason to make decisions like humans. Some of the major areas where pattern recognition finds applications are Cognitive communications, Computer vision, Satellite imaging and satellite image analysis, Medical Image Analysis, Printed character recognition, Handwriting recognition, Geo-informatic systems, Speech recognition, Biometric identification, Document classification, Internet search engines and Credit scoring etc.

The pattern recognition scheme [1],[8] usually adopts either statistical or syntactic (structural) approach. The former is based on statistical characterizations of patterns, assuming that the patterns are generated by a probabilistic system while the latter is based on the structural interrelationships of features. Minimum distance classifiers and correlators are known as template matching techniques and are related to statistical approach. This paper deals with one such pattern matching technique based on the use of permutation invariant Rao Alaka Shift Transform (RAST). RAST transforms raw data of spatial domain into highly correlated data of the transformed domain. Hence image patterns are more effectively recognized in the transformed domain by adopting appropriate template matching techniques.

A character of an alphabet is written in a specific manner and hence the digital image due to a character can be considered as the pattern associated with it. When characters are printed in different fonts, the digital pattern of a particular character in one font bears some common features with patterns of the same character in other fonts. Hence, the digital images due to same character in various fonts can be treated as the members of a class of that character. In other words, one can conceive as many character classes as there are characters in the alphabet. Similarly, taking a printed character in one font and skewing it clockwise or anticlockwise by varying degrees one can generate a different class of patterns pertaining to the character. This paper explains a method where digital image patterns due to various printed characters of English alpha numerals in several fonts are recognized and classified.

2. RAS TRANSFORM AND ITS PI PROPERTIES

RAS Transform with its Permutation Invariance (PI) properties has been used for the sake of binary and gray level image processing and coding of Multimedia digital data[2],[3],[9] at Centre of Research in Electronics and Information Technology (CREIT), Hyderabad. RAST is formulated by of AVN Rao and Alakanandana and hence it is monikered as Rao Alaka shift Transform (RAST). RAST is a simple transform which uses only relational and shift operations involving very less mathematical computation and also a fast executing Transform. RAST transforms a number sequence x(n) of length N, equal to any power of 2, into highly correlated sequence X(k) of same length called its Transform. An encryption key E(r) can be generated during forward transformation which helps for inverse transformation. RAST introduces isomorphism that maps a domain set consisting of a sequence of symbols, its dyadic permutations, their graphical inverses and all their cyclic shift permutations onto a range set consisting of a unique sequence of symbols, called its RAS Transform and its encryption key.

This map is of one-to-one and onto correspondence and an inverse map also exists and hence viewed as a transform. This map ensures the invariance property under such permutations. RAS Transform is applicable for the binary, decimal and both positive & negative numbers with any base. It is most generalized and can as well be applied to any set of symbols or numbers with known hierarchy. The permutation invariance and Second-Order Self invariance property of RAST make coding of digital data with RAST equally attractive like other transforms, such as FFT, GRT etc, which are used in several applications of Digital Image Processing, speaker/speech verification [4], Encryption [5]. RAST is also used for Error Detection [5] and for Pattern Recognition [2].

Algorithm for implementation of RAS Transform is nonlinear and recursive. There are two methods in which RAST can be implemented based on the order of hierarchy chosen, namely, ascending mode of implementation (AMI) and descending mode of implementation (DMI). Let x(n) be a number sequence such that its length N=2m ; m > 0, is satisfied and its RAST is denoted as X(k). The algorithm is implemented recursively in 'm' iterations. It generates an encryption key of 12 bits for an eight point input sequence. The encryption key is used for inverse transformation.

Example: If the given input binary and non-binary sequences are x1(n)={10110011} and x2(n)={45731892} then after application of RAST, DMI, they would be of the form X1(k) ={11101100} with E1(r) ={000001110000} and X2(k) ={98437521} with E2(r) ={011010100001}.

2.1. PI properties of RAS Transform

The RAS Transform exhibits three types of permutation invariance properties. They are i) Cyclic shift invariance property ii) Graphical inverse invariance property and iii) Dyadic shift invariance property. The first two properties together are known as translation invariant property. Let us consider a number sequence x(n)={7,5,8,3,2,6,1,9}. Using this sequence one can generate seven more cyclic shift versions. When the sequence is cyclic shifted eight times, the sequence x(n) is again obtained. It is found that all the eight sequences have the same RAST, X(k), as {9, 8, 7, 6, 5, 2, 3, 1} but different E(r)s..

For x(n)={7,5,8,3,2,6,1,9} the graphic inverse is x-1(n) ={9,1,6,2,3,8,5,7}. Using this sequence one can generate seven more cyclic shifted versions. Similar to x(n), x-1(n) also has eight mutually independent sequences. Again it is verified that all the eight sequences, representing the cyclic permutation class of x-1(n), have the same RAST, X(k), but different encryption keys.

The term dyad refers to a group of two and the dyadic shift operation is transposition of two blocks of elements in a sequence. For instance let x(n)={7,5,8,3,2,6,1,9} and on exchange of its first half with the second half it results in a sequence {2,6,1,9,7,5,8,3}. This process is continued by further dividing one half into two more parts and so on, one can get 8 sequences. Again, it is verified that all these sequence have the same X (k) but different E(r)s. Further, taking graphical inverses and then cyclic permutations for these two dyadic sequences, then in all, there would be 48 mutually independent permutations, all derived from x(n). Again, quite fascinatingly, it is found that for all these 48 sequences, the RAS transformed output is a single sequence, X (k), while the encryption keys, E(r)s, are different. These 48 sequences form the permutation class of x(n). This property of RAST enables it to exhibit rotational invariance as exhibited by other transforms [2].

If the number sequence x(n) is binary in nature, then it is possible for some of them to have less number of mutually independent sequences than expected 48. At the same time, it is found that there could be one or more binary sequences x1(n) and x2(n) and so on which do not fall under the cyclic shift permutations of a given binary sequence x(n) and yet may have the same RAST, X(k). However, without any exception, the encryption key codes, E(r)s, will be different.

2.2. Second order Invariance property of RAS Transform

This property says that once RAST of a number (symbol) sequence is obtained, any number of repeated applications of RAST on the transformed number sequence does not give riseto any further change. i.e) Application of RAST on a number sequence introduces a definite order in the sequence and hence RAST can also be called sequence organizing transform. To be more precise, given any arbitrary sequence of appropriate length, the second and higher order RASTs would be invariant. Thus, RAST exhibits the second-order self invariance property. This particular property is recognized as very useful in error detection/correction applications.

3. METHODOLOGY

The methodology adopted here is essentially based on a template matching approach. This type of pattern matching system generally comprises of two phases namely database generation phase, also known as training phase and testing phase. The first step is data acquisition and data conditioning. Raw data corresponding to the object of recognition is acquired and pre-processed to be in suitable form for further treatment. Next, the data corresponding to different patterns of several classes is stored in the data base. Subsequently, a template for each of the pattern classes is generated, based on the available patterns of various classes, and stored.

In the second stage of testing, a test pattern is used in combination with the template of each one of the possible classes and an appropriate metric is evaluated. Based on the metric and the nature of its value a decision is made regarding the class to which the test pattern belongs. A set of results are generated by repeatedly testing with meaningful number of test patterns and then classifying them. The results are analyzed to estimate the efficacy of the pattern identification system.

In this paper, the black and white digital images pertaining to 20 characters, randomly chosen 10 letters (A, B, E, G, K, O, M, W, Z) of English alphabet and 0-9 numerals, printed in five different fonts, a total of 100 character patterns are obtained. The binary matrix representing the digital image of each character is subjected to the permutation invariant RAS transform and its transformed domain version is also computed. One template or archetype each is generated for the characters in both the spatial and transformed domains.

A digital image is essentially represented by a matrix of numbers corresponding to brightness of image pixels. Hence, the simple method that is adopted here is to compute the average of matrices pertaining to one character in various fonts and call it the template for that character class. Five fonts are considered in our study. The fonts used are Times New Roman (TNR), Bookman Old Style (BOS), Baskerville

Old Face (BOF), Century (CEN) and Verdana (VER). A single set of templates is created jointly for both the 'Normal' and 'Bold' fonts. Euclidean distance is the metric used to decide the match with the template. The class membership of unknown character pattern is decided by its closeness or minimum Euclidean distance to the class template. Hence the problem of pattern recognition is reduced to computing the distance metric. The formulae used [7] for this purpose are given below. Template of a character class is generated by computing the mean of the matrices pertaining to the digital images of the character, under consideration, in various fonts. Equation (1) shows how the mean matrix or class template is evaluated.

$$\mathbf{m}_j = \frac{1}{N_j} \sum_{x \in w_j} \mathbf{x}_j \qquad j = 1,2.......W \quad \text{--(1)}$$

Here, x is the unknown pattern matrix, Nj is the number of pattern matrices from class wj and summation is taken over these matrices. In our case Nj is 5. Total number of classes considered are W=20, as twenty characters are used for this study. The equation for computing the distance measure is given by (2). The absolute distance between the template and the test pattern is D(x).

$$D_j(\mathbf{x}) = \|\mathbf{x} - \mathbf{m}_j\|$$

$$j = 1,2.......W \qquad \text{------(2)}$$

where $\| \mathbf{a} \| = (\mathbf{a}^T \mathbf{a})^{\frac{1}{2}}$ is the Euclidean norm. x is assigned to class wj if Dj(x) is the smallest distance.

Matching between the test pattern and the template also can be accomplished by computing the cross correlation function (XCF) between the functions corresponding to these two images, as given in (3). For ease of computation, the 2-D matrix corresponding to a character pattern is read into 1-D matrix and then the XCF is computed. The template of whichever character class gives the maximum value for the computed XCF, the pattern is assigned to that character class.

The correlation of two functions f(x,y) and h(x,y) is defined as

$$f(x,y) \circ h(x,y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f^*(m,n) h(x+m, y+n) \qquad (3)$$

It is observed that the values obtained for match with template used in transformed domain and computation of XCF are very close with each other and these are much greater than the value of match obtained in the spatial domain. This observation validates the statement that the data in the transformed domain is highly correlated compared to the one available in the spatial or time domain.

The RAS transform has permutation invariance as one of its important properties. To take the advantage of this property, the transform is tested to identify the characters even when they are tilted or slanted by varying degrees, up to 250, both in clockwise and anticlockwise directions. The characters in one font are only used for this study. The non skewed version of the character image is taken as the template in both the spatial domain and in the transformed domain. The skewed versions are all taken as the test patterns to be recognized and classified. The results are presented in the next section and discussed.

4. RESULTS AND DISCUSSIONS

The digital images obtained for various characters of 'Bookman Old Style' font in spatial domain are shown in figure-1. Similarly, images are acquired for all 20 characters in other five different fonts. To give the idea of variation from font to font five characters in five different fonts are shown, in figure-2. Pattern (character) recognition is attempted in the spatial domain by template matching. A template for that specific character classis generated based on eqn.(1).

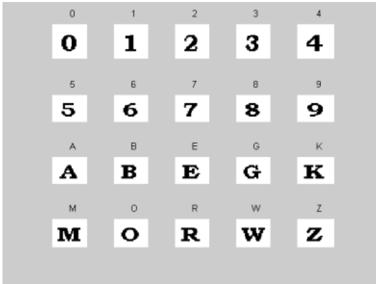

Figure1: Spatial domain digital images of characters in 'Bookman Old Style' font

The decision of a character recognized is done based on the computed minimum Euclidean distance as per the eqn.(2). If all characters are correctly identified, irrespective of the font, then the character recognition capability is 100% and if only 10 out of 20 characters are correctly recognized then it is only 50%. Accordingly, the values are found by testing the entire 100 patterns n spatial domain

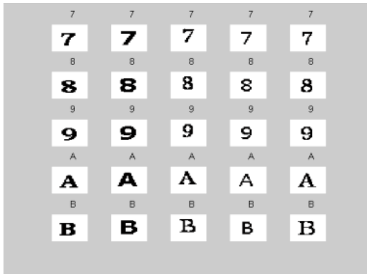

Figure 2: Spatial domain digital images of characters in 'Bookman Old Style' font

In this process no separate parameters, connected with any of the features of the images or patterns, are computed. Instead, the mean matrix, representing the image as a whole, is taken as template.

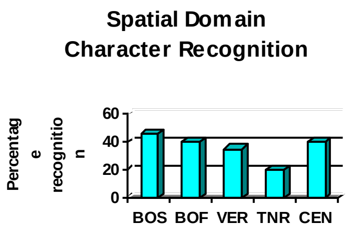

Figure-3: Results of Character Recognition in Spatial Domain

The bar diagrams in figure-3 shows the pattern recognition in percentage. The diagram shows the various percentages obtained for different fonts based on the number of matches obtained during the testing phase. It is seen that the Character recognition, in the spatial domain, between the template and the test pattern is quite low. The over all percentage recognition of a character is mere 36.

Subjecting characters in 'Bookman Old Style' font to permutation invariant RAS transform gives the images, in the transformed domain, as shown in Figure-4. As stated earlier, these images, in transform domain, represent highly correlated data pertaining to the character and are useful for pattern recognition.

RAS Transform is applied to images of all 20 characters in five different fonts and the database of test patterns is generated also in the transform domain. A template, in transformed domain also for each of the character class, is generated by taking the mean matrix from the transform domain representations of the same character taken in five different fonts.

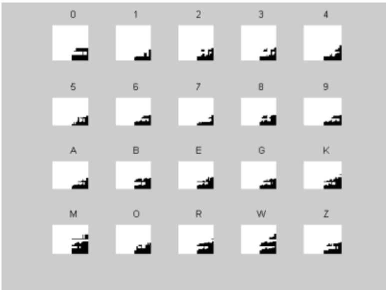

Figure 4: Transformed domain Digital images of Characters in 'Bookman Old Style' font and used as templates for skewed pattern recognition.

The same procedure is adopted for all the 20 characters used for this study. These templates are shown in figure-5. Again template match in the transformed domain is adopted and the values of match are found by testing all the 100 patterns.

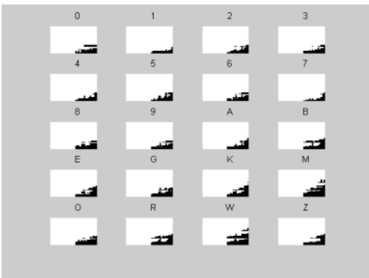

Figure.5: Transformed domain Templates of Characters used for pattern classification of characters in five different fonts

The results of character recognition in transformed domain are shown in figure-6. The bar diagrams here, show the pattern recognition in percentage obtained for five different fonts based on the number of matches obtained during the testing phase.

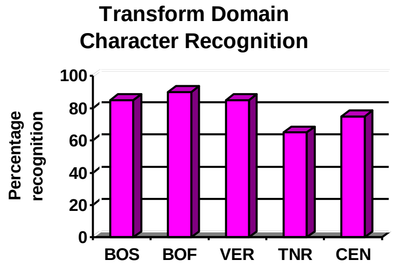

Figure 6: Results of Character Recognition in Transform Domain

The percentage character recognition has considerably increased though it has not become 100%. Based on the fonts chosen and they are either bold or normal, the percentage recognition varied from font to font when several combinations are tested. The overall percentage recognition is 80. This is more than double the corresponding value in spatial domain.

The permutation invariant property of RAS transform is expected to be helpful in pattern matching of rotated or skewed characters by varying degrees in both clockwise and anticlockwise directions. To verify this, the characters are skewed to left and right in steps of 5^0, on either side, up to 25^0 and both in spatial domain and their RAS transformed versions also obtained for generating the database of test patterns.

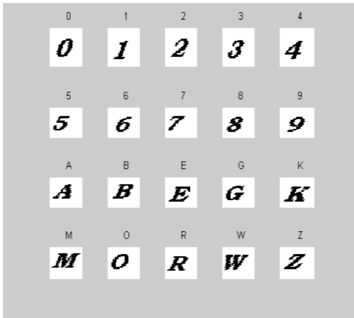

Figure.7: Spatial domain images of skewed Characters by 25^0 clockwise in 'Bookman Old Style' font.

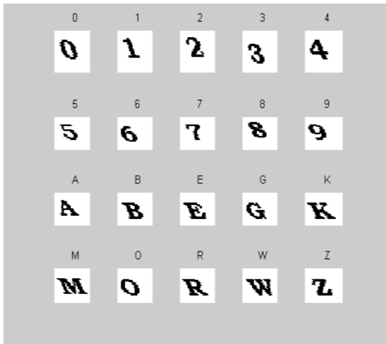

Figure8: Spatial domain images of skewed Characters by 25^0 anticlockwise in 'Bookman Old Stvle' font.

Figure-7 and Figure-8 show the skewed (by 25^0) patterns of characters in 'Bookman Old Style' in clockwise and anticlockwise directions respectively.

The direct patterns corresponding to normal characters in that font, shown in figure-1, are taken as templates in the spatial domain. Similarly, the simple patterns corresponding to RAS transformed versions, shown in figure-4, are taken as templates in the transformed domain. Again, template matching is performed and the results of skewed character recognition are shown in the fig.9, a bar diagram.

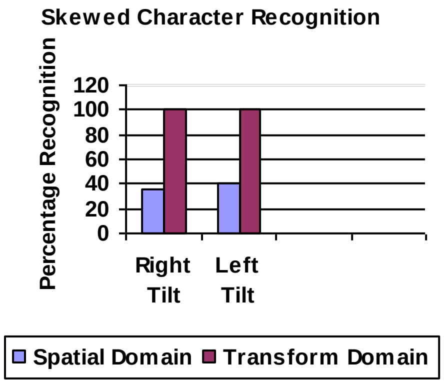

Figure-9: Results of Skewed Character Recognition

It is evident from the results that as per expectation, the permutation invariant Rao Alaka Shift transform has given excellent results of recognition for rotated patterns. It is found in this study, that up to 45^0 of tilt, the characters are recognized with 100% recognition accuracy and rotation invariance in pattern recognition is accomplished 100% within this limit. Subsequently, the recognition capability reduced as the angle of skew increased towards 90 degrees. It is also observed that there has been again 100% recognition when the test pattern is the graphical inverse of any one of the already well recognized patterns. This particular capability of RAS transform is reported for component recognition in a factory automation [2] problem for autonomous classification and segregation of machine parts in a production shop.

5. CONCLUSIONS

The high correlation of transformed data, using permutation invariant RAS transform, has enhanced the pattern recognition and classification capability from less than 40% level in spatial domain to 80% in the transformed domain.

The requirement of memory and mathematical complexity is reduced as there was no need to compute feature parameters to generate templates by using RAS transform. Permutation invariant properties have contributed effectively to obtain 100% pattern recognition and classification of tilted characters.

Combining neural computing, a well recognized tool for pattern recognition, with the PI transforms is likely to open new avenues for achieving complete rotational invariance and accomplishing better pattern recognition and classifications.

6. ACKNOWLEDGMENTS

The authors wish to thank the Managing Director, Centre of Research in Electronics and Information Technology, Hyderabad for permitting us to interact with the scientists of the centre. The authors also thank their respective principals for permitting them to work on this important area of Image processing and Pattern recognition.

7. REFERENCES

[1] S. Furui, "An overview of speaker recognition technology", *ESCA Workshop on Automatic Speaker Recognition, Identification and Verification*, pp. 1-9, 199.

[2] Narasimha Rao. A.V, Srinivasa Rao K, M. Srinu Naik and Alakanandana.A " The Role of Permutation Invariant RAS Transform in the Machine Vision Based Object Identification", Proceedings of 4th National Conference on 'Applications of Emerging Technologies', Adhiyamman College of Engineering, Hosur, 4-5, April 2007.

[3] A.V. Narasimha Rao, K. Srinivasa Rao & A. Alakanandana "Permutation Invariant Coding of Digital Data", Proceedings of IETE Golden Jubilee Symposium on Convergence of Technologies, IETE Hyderabad Centre, Hyderabad, 11-13th July, 2003.

[4] Narasimha Rao A.V., K. Srinivasa Rao & VSL Sowmya "Speaker Verification using Permutation Invariant Transform", Proceedings of National Conference on Recent Trends in Electronics & Communication, SUS College of Engineering & Technology, Tangori, Punjab, 10-11th April, 2008.

[5] Narasimha Rao A.V., Soundara Rajan K & Srinivasa Rao K. "Integrating Error Detection with Data Encryption Algorithm using Permutation Invariant Rao Alaka Shift Transform" ARPN Journal of Engineering and Applied Sciences, ISSN 1819-6608, Vol. 3, No. 5, October 2008.

[6] M.D Wagh and S.C. Kanetkar, ' A class of Translation Invariant Transforms', IEEE transactions on Acoustics, Speech, Signal Processing, pp. 203-205, April 1977.

[7] Rafael C. Gonzalez & Richard E. Woods, Digital Image Processing, second edition, pp 720-721, Pearson Education Inc., 2002.

[8] Rajan E.G,A.v.Narasimha Rao, "Role of GRT in Machine Vision in Factory Automation",Proceedings of National Conference on Intelligent Manufacturing systems,Feb,1998,Coimbatore

[9] A.V.Narasimha Rao,"Fast Computing Transform for image Processing" Proceedings of XVII National Convention of Electronics & telecommunication Engineers, Hyderabad, October 2001.