

Dual Vector Features for Rotation-Invariant Handwritten Character Recognition

Dr. Priyadarshan Dhabe¹, Aditya Bodhankar², Parth Sheth³, Srushti Shevate⁴ and Diksha Prasad⁵

¹⁻⁵Vishwakarma Institute of Technology, Pune, India

Email: priyadarshan.dhabe@vit.edu, aditya.bodhankar21@vit.edu, parth.sheth20@vit.edu, srushti.shevate20@vit.edu, diksha.prasad21@vit.edu

Abstract— In the field of image processing, invariance refers to the properties of an image that remain unchanged or exhibit minimal differences under transformations like rotation, scaling, or blurring. The primary focus of this work is to overcome the challenge of recognizing handwritten characters from various rotation angles. To address this, we propose a novel approach that utilizes modified quad-vector features proposed in [7], enabling the accurate identification of pixels associated with specific rings. Our main objective is to achieve efficiency and the rotation invariance too. To accomplish this, we employ only two vectors to extract features rather than four as suggested in [7]. Hence the name dual vector features are given. The extracted dual vector features establish a framework for extracting rotation-invariant features (RIFS) from black-and-white images. The proposed approach is found 1.13 times faster than the original method [7] and thus efficient and recommended for Handwritten Character Recognition (HCR) systems.

Index Terms— Handwritten character recognition, Rotation Invariant Features, Ring features, Rotation Invariance.

I. INTRODUCTION

HCR is essential in various applications, including document processing and postal sorting systems. However, accurately recognizing characters that undergo different rotations or variations in orientation remains challenging. Invariance to such transformations is crucial for robust character recognition systems that can operate effectively in real-world scenarios. The incorporation of rotation-invariant features in character recognition is aimed at improving accuracy and adaptability. Traditional algorithms often struggle with rotated or skewed characters due to significant changes in strokes and shapes caused by orientation variations. By introducing rotational invariant characteristics into the recognition process, HCR systems can achieve improved accuracy and resilience. This research paper addresses the rotation invariance problem in HCR by proposing modified quad-vector features defined in [7]. It discusses a unique method for recognizing and capturing pixel data related to individual character rings in handwritten characters. This approach ensures a precise representation and characterization of the characters, regardless of their rotation. The main focus is on achieving computational efficiency by efficiently calculating distances. The method uses the Euclidean distance to measure the distance between the object pixels and two vertices of the image frame. The ultimate objective is to optimize the distance calculation process and improve the efficiency and effectiveness of rotation-invariant HCR (Handwritten Character Recognition). The subsequent sections of this paper will detail the methodology, experiments, and results, providing insights

into the effectiveness and potential applications of the proposed dual vector features, which is modification to quad-vector features approach [7] for rotation-invariant HCR. By addressing the rotation invariance challenge, this research aims to contribute to the faster, robust and accurate character recognition systems, with implications across various domains such as document analysis, optical character recognition, and pattern recognition.

The format of this document is as follows. Section II demonstrates the literature survey. Section III describes the drawbacks of the quad vector features mentioned in [7]. The proposed dual vector feature method is discussed in section IV. Comparison of time complexities of proposed and existing approaches is done in section V. Conclusions are withdrawn in section VI and references are cited at the end.

II. LITERATURE REVIEW

Din-Chang Tseng, Hung-Pin Chiu, Jen-Chieh Cheng [1] address the recognition challenges associated with handwritten Chinese characters through a novel approach based on fuzzy min-max (FMM) neural networks. The paper emphasizes the limitations of traditional methods in handling variations and deformations, leading to reduced accuracy. To overcome this, the authors propose utilizing FMM neural networks, known for their ability to handle uncertainties and variations in data. The proposed methodology involves preprocessing steps such as image normalization and feature extraction, with the extracted features serving as input for the FMM neural network model. The research includes experimental evaluations on a suitable dataset, utilizing evaluation metrics to assess recognition performance. The results demonstrate significant improvements in accuracy and robustness compared to traditional methods. This research contributes to the advancement of handwritten Chinese character recognition, offering a promising approach for applications in document processing, optical character recognition, and pattern recognition.

Hung-Pin Chiu and Din-Chang Tseng [2] proposed an innovative recognition system that utilizes a fuzzy neural network to identify handwritten Chinese characters on maps. The focus of their research is on reading handwritten Chinese characters with arbitrary location, scale, and orientation on drawings or maps. The first step involves normalizing binary character patterns to achieve scale and translation invariance. Subsequently, simple and rotation-invariant ring data are extracted from thinned or non-thin normalized characters. The study introduces an invariant recognition system that relies on fundamental invariant ring-data features and a powerful fuzzy min-max neural network, specifically designed for recognizing handwritten Chinese characters on Chinese land-register maps.

Meneganti, Saviello, and Tagliaferri [3] introduced a novel learning algorithm for the Simpson's fuzzy min-max neural network. This new algorithm addresses certain drawbacks of the Simpson's model, such as the absence of sensitivity parameters or thresholds that could restrict the dimension of hyperboxes. By utilizing this algorithm, the network's performance is significantly improved, as the classification outcome becomes independent of the order in which patterns are presented in the training set. Additionally, the classification error in the training set does not increase with each step. A comparative analysis with other fuzzy neural networks from existing literature demonstrates the advantageous nature of this new neural model in handling classification problems.

In their work, Gabrys and Bargiela [4] introduce the generic fuzzy min-max (GFMM) neural network, which serves as an extended version of Simpson's fuzzy min-max clustering and classification techniques. The GFMM method combines supervised and unsupervised learning in a unified training process, making it applicable for pure clustering, pure classification, or hybrid clustering classification tasks. A notable feature of the GFMM neural network is its ability to effectively distinguish decision boundaries between classes, resulting in improved accuracy and robustness when dealing with intricate data patterns. The paper provides a comprehensive explanation of the GFMM approach, accompanied by practical examples and a comparative analysis with Simpson's methods. Additionally, the researchers showcase the application of GFMM in detecting water distribution system leaks. The GFMM method has undergone efficiency enhancements, broadening its potential for various applications, including handwritten character recognition. By leveraging both supervised and unsupervised learning, the GFMM neural network exhibits promising results in recognizing handwritten characters with different styles and orientations. Overall, the GFMM approach represents a significant advancement in clustering, classification, and handwritten character recognition, opening up exciting possibilities for image processing applications.

U.V. Kulkarni, T.R. Sontakke, and G.D. Randale [5] proposed the fuzzy hyperline segment neural network (FHLSNN) as an innovative approach for rotation-invariant handwritten character recognition (HCR). The FHLSNN algorithm was designed to address the challenge of recognizing handwritten characters from various angles by incorporating hyperline segments, which effectively handle rotation invariance. The study compared the FHLSNN algorithm with two other neural network approaches, namely Patrick Simpson's fuzzy min-max neural network and Kwan and Cai's fuzzy neural network, in a supervised environment. The comparative analysis

demonstrated that the FHLSNN outperforms these alternatives in terms of generalization, recall time per pattern, and training time, establishing its superiority in HCR tasks. To further enhance recognition accuracy, the researchers applied linear moment normalization to the handwritten characters, ensuring scale and translation invariance. This normalization technique enables the FHLSNN to accurately recognize characters regardless of their size or position, making it a robust and reliable solution for practical handwritten character recognition systems. Overall, the FHLSNN represents a promising advancement in the field of HCR, offering an efficient and effective method for achieving rotation invariance and improving the overall performance of handwritten character recognition systems.

Kulkarni, Uday Vasantrao and Sontakke T. R, [6] addresses the issue of recognizing rotation-invariant handwritten Devanagiri numerals. The goal of this thesis is to study existing fuzzy neural networks and to propose new architectures with their learning algorithms, which can be used in pattern classification and clustering applications. Five novel algorithms are offered in this thesis as intriguing expansions to several of the current ones. The first approach is an expansion of a fuzzy neural network created by Kwan and Car that uses similarity metrics to build neurons in the output layer in an unsupervised environment. The fuzzy hyperline segment neural network, the second suggested design, uses fuzzy sets as pattern classes, with each fuzzy set being the union of fuzzy set hyperlink segments. The next design, dubbed fuzzy hyperline segment clustering neural network, may quickly pick up on poorly defined nonlinear cluster boundaries. The fuzzy hypersphere neural network classifier, which uses fuzzy sets as pattern classes and each fuzzy set as a union of fuzzy set hyperspheres, is another unique design that has been proposed. Finally, a generalized fuzzy hypersphere neural network that combines supervised and unsupervised learning is developed. It is an expansion of the fuzzy hypersphere neural network. On a Pentium III computer running at 733 MHz, the suggested methods have been evaluated and implemented using the MATLAB software environment.

P.S. Dhabe, S.R. Bakki, U.V. Kulkarni, and T.R. Sontakke [7] proposed a new approach to retain rotational invariant features (RIFs) for monochrome images. The approach is modified to get RIFs consistent for all angles, minimizing variations in RIFs of a picture when rotated at angles other than multiples of 90. The latter is the one that this research emphasizes because the classifier's choice is purely based on invariant traits. The suggested method with modification produces features at all feasible rotational angles that are significantly more rotation invariant than others. Comparing the suggested method's discrimination power to ring data and Zernike features, it performs better. This method is found to be superior to Zernike features in terms of both rotation invariance and cost. With or without taking into account an image's boundaries, this approach provides better rotation invariance. P.M. Patil, U.V. Kulkarni, and T.R. Sontakke [8] extended the fuzzy hyperline segment neural network proposed by Kulkarni, Sontakke, and Randale by proposing a general fuzzy hyperline segment neural network (GFHIS) and its learning algorithm. It combines clustering and classification, enabling it to perform pure clustering, pure classification, or hybrid clustering/classification tasks. Notably, it demonstrates efficient hyperbox adaptation, resulting in reduced hyperboxes and improved recognition rates. The algorithm's rapid training process and simple operations contribute to its high efficiency. In the context of handwritten character recognition, the GFMM's ability to handle fuzzy input patterns and decision boundary identification make it a promising tool for accurately recognizing and classifying handwritten characters. However, it is essential to provide representative training data for optimal performance, as the algorithm's efficiency may be affected by significant differences between training and test data.

M. Hanmandlu, K.R. Murali Mohan, Sourav Chakraborty, Sumeer Goyal, and D.Roy Choudhury [9] proposed the box method for feature extraction, which is a revolutionary strategy for the recognition of handwritten characters. This paper makes an effort to identify unrestricted or free handwritten characters. This study seeks to create a precise system that targets 100% facial recognition in a range of sizes, shapes, and fonts. The authors of this study present a way allowing handwritten numerals (a-z), letters (A-Z), and numbers (0-9) to be recognized.

P.M. Patil, P.S. Dhabe, U.V. Kulkarni and T.R. Sontakke [10] modified membership function of fuzzy hyperline segment neural network (FHLSNN) proposed by U.V.Kulkarni and T.R.Sontakke. In Handwritten Character Recognition (HCR), the paper proposes a modified MFHLSNN algorithm with rotation-invariant ring features and linear moment normalization. The algorithm is trained on a dataset of handwritten Devanagiri numerals, achieving superior recognition rates and faster training and recall times compared to other algorithms, making it reliable for real-time HCR features. The improved membership function is shown to be superior to the function developed in the study, which at first awarded lower values to patterns near the hyperline segment but far from its end points. Using two splits of the FISHER IRIS data, the improved MFHLSNN's performance is assessed, and it is found to perform better than FHLSNN in terms of recognition rates.

Abhijeet V. Nandedkar, Kishore Venishetti and Ajendra Kumar Rathod [11] developed an invariant to translation, rotation, and scale character recognition system. The feature extraction and recognition processes make up the

system's two primary parts. RTSI (Rotation, Translation, and Scale Invariant) characteristics are used for feature extraction. This feature vector's key benefit is that character normalization is not necessary. Compared to alternative approaches, it is quite easy to implement these functionalities. 26 uppercase English capital letters typed in a variety of fonts, including Ariel Unicode, Ariel Narrow, Microsoft scan serif, and handwritten characters, are used to test character recognition algorithms. An excellent invariance under translation, rotation, and scaling is achieved by the proposed invariant character recognition (ICR). The primary contribution of this study is the creation and application of new features on character recognition techniques for enhancing performance. This system is reliable, quick in computation, and simple to use.

Pradeep M. Patil and Manish P. Deshmukh [12] described a modification of the General Fuzzy Hyperline Segment Neural Network (GFHLSNN) introduced by Patil, Kulkarni, and Sontakke is the Modular General Fuzzy Hyperline Segment Neural Network (MGFHLSSNN) with its learning algorithm. The rotation-invariant handwritten character recognition method MGFHLSSNN, an extension of GFHLSNN, was presented in this study. The proposed algorithm is applicable to pure clustering, pure classification, and hybrid clustering and classification.

III. DRAWBACKS OF QUAD-VECTOR FEATURES

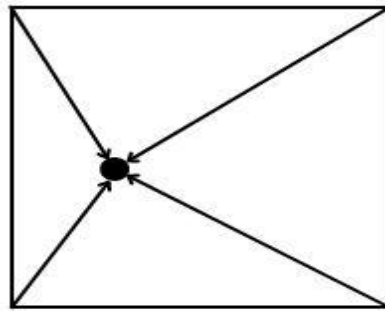


Figure 1: Quad vectors

One drawback of quad vectors as seen in Fig.1, is the requirement to calculate four Euclidean distances between vectors and a pixel. This entails computing the four distances per image pixel, as shown in Fig. 1. This fact increases time complexity of the algorithm, leading to a potential impact on runtime.

The need to compute four Euclidean distances per pixel, makes working with quad vectors [7] time-consuming and computationally intensive. The increased number of computations can significantly slow down the overall process of feature extraction. This effect becomes more pronounced when dealing with large datasets or large image sizes.

IV. PROPOSED DUAL VECTOR FEATURES

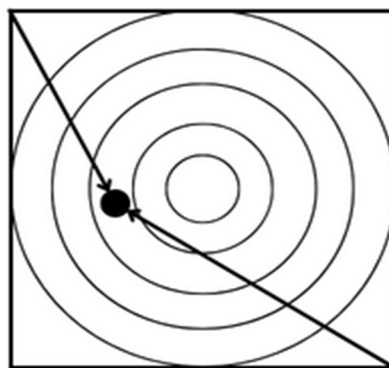


Figure 2: Dual Vectors.

The digital image depicted in Fig. 2, is of size $N \times N$, containing N rows and N columns of pixels, with white pixels (object pixels) of the character distributed in concentric rings. To locate individual white pixel, two vectors are drawn from any two opposite corners of the image to a given pixel as illustrated in Fig. 2.

If the location of the pixel has the coordinates (x_1, y_1) and if the vector's tail is at (x_2, y_2) , its magnitude m , of this vector is given by the Euclidean distance as shown in (1).

$$m = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

and the angle of vector, θ , is given as in (2),

$$\theta = \tan^{-1} \left(\frac{y_2 - y_1}{x_2 - x_1} \right) \quad (2)$$

The term "*Dual vectors*" is defined as the two vectors which are drawn from any two vertices of an image that are opposite, to a white pixel of the character, as shown in Fig. 2. The magnitudes of the two vectors for a pixel are added up to form the dual vector sum also called Qsum. If we assume that v_1, v_2 represents the vector magnitudes. These vectors are drawn from two points on the image to the current pixel (x, y) . Based on this we can calculate the Qsum of a character image $f(x, y)$ as shown in (3).

$$Q_{sum} = \sum_{x=1}^N \sum_{y=1}^N (v_1 + v_2) f(x, y) \quad (3)$$

Qsum is calculated for all the pixels located in the particular ring. For example, if there are four pixels in the ring, then Qsum for four pixels will be calculated like $Qsum_1, Qsum_2, Qsum_3, Qsum_4$. At last 'Total Qsum' is calculated which will be the sum of all Qsums in the particular ring, which can be described as in (4).

$$Total \ Qsum = \sum_{i=1}^n Qsum_i \quad (4)$$

TABLE I. COMPARISON OF EXECUTION TIMES.

Run	Dual vectors	Quad Vectors
1	646 ms	680 ms
2	663 ms	896 ms
3	698 ms	792 ms
4	755 ms	723 ms
5	725 ms	883 ms
6	658 ms	697 ms
7	701 ms	796 ms
8	762 ms	887 ms
9	691 ms	815 ms
10	760 ms	851 ms
11	807 ms	883 ms
12	712 ms	721 ms
13	653 ms	654 ms
14	726 ms	685 ms
15	808 ms	811 ms
16	770 ms	863 ms

17	802 ms	889 ms
18	745 ms	802 ms
19	700 ms	902 ms
20	737 ms	1030 ms

Table 1, shows the comparison of feature extraction time needed for both the methods for a single image. Since, the measured execution time varies per run, we considered time for 20 runs and its average.

The graph in Fig. 3 shows the comparison of execution time of dual vectors and Quad vectors, taken for 20 runs. The average time taken by proposed dual vectors is 706.8 milliseconds and for Quad vectors [7] is 803.55 milliseconds. It can be clearly seen that the time required for dual vectors is less if we compare it with the time required for the Quad vectors [7]. The proposed approach is at least 1.13 times faster than the original method [7].

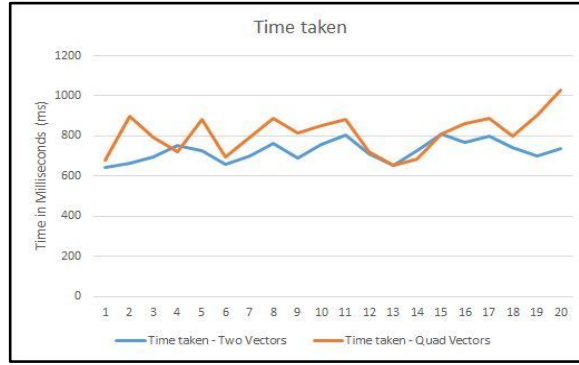


Figure 3: Comparison graph of execution time of both the methods

V. COMPARISON OF TIME COMPLEXITIES

In the quad vector features [7], we need to calculate the norm or length of 4 vectors v_1 , v_2 , v_3 and v_4 , from each corner of the image frame to a single pixel, see Fig.1. Thus, the time need to calculate these features of an image with size $N \times N$ can be calculated as follows.

1. Number of distances need to be calculated for a single pixel is

$$T_1 = 4.K$$

$K > 0$ constant time needed for calculating one distance

2. Total time $T(N)$ need for processing all the pixels in image can be computed as

$$\begin{aligned} T(N) &= N \times N \times 4.K \\ T(N) &= 4KN^2 \end{aligned} \quad (5)$$

Similarly, from (5), one can find the time required for the proposed dual vector features as shown in (6).

$$T(N) = 2KN^2 \quad (6)$$

Since, only two distances need to be calculated for the dual vector method. The time complexity of both approaches is $O(N^2)$ but the proposed dual vector method is found to be twice faster than the original one. This conclusion is obvious from comparison of (5) and (6) and also verified from experimentation given in Table 1.

VI. CONCLUSIONS

In this paper, we proposed a dual vector method to rotation invariant feature extraction, which is a modified and/or optimized version of the quad vector method proposed in [7]. The average time taken by proposed dual vectors is 706.8 milliseconds and for Quad vectors is 803.55 milliseconds, for feature extraction of a single image. Thus, the proposed method is, on an average, 1.13 times faster than the original method [7]. The time complexity of both the approaches is exactly same but the proposed method has reduced the constant multiplied with the term KN^2

from 4 to 2. Thus, theoretically, the proposed method of dual vector is twice faster than the original one [8] and thus recommended strongly for rotation invariant feature extraction for HCR for practical applications.

REFERENCES

- [1] Din-Chang Tseng, Hung-Pin Chin, Jen-Chieh Cheng, Invariant handwritten Chinese character recognition using fuzzy ring data, *Image and Vision Computing*, Volume 14, Issue 9, 1996.
- [2] Hung-Pin Chiu, Din-Chang Tseng, Invariant handwritten Chinese character recognition using fuzzy min-max neural networks, *Pattern Recognition Letters*, Volume 18, Issue 5, 1997, Pages 481-491, ISSN 0167-8655
- [3] M. Meneganti, F. S. Saviello and R. Tagliaferri, "Fuzzy neural networks for classification and detection of anomalies," in *IEEE Transactions on Neural Networks*, vol. 9, no. 5, pp. 848-861, Sept. 1998.
- [4] B. Gabrys and A. Bargiela, "General fuzzy min-max neural network for clustering and classification," in *IEEE Transactions on Neural Networks*, vol. 11, no. 3, pp. 769-783, May 2000.
- [5] U. V. Kulkarni, T. R. Sontakke and G. D. Randale, "Fuzzy hyperline segment neural network for rotation invariant handwritten character recognition," *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222)*, Washington, DC, USA, 2001, pp. 2918-2923 vol.4.
- [6] Kulkarni Uday vasantrao, "Rotation invariant handwritten character recognition a fuzzy neural approach", Ph.D thesis, Swami Ramanand Teerth Marathwada University, Nanded, Maharashtra, India, 2001.
- [7] P. S. Dhabe, S. R. Bakki, U. V. Kulkarni and T. R. Sontakke, "Rotation invariant features: a method, problems and possible solutions," *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02'*.
- [8] P. M. Patil, U. V. Kulkarni and T. R. Sontakke, "General fuzzy hyperline segment neural network," *IEEE International Conference on Systems, Man and Cybernetics, Yasmine Hammamet, Tunisia, 2002*, pp. 6 pp. vol.4.
- [9] M. Hanmandlu, K.R. Murali Mohan, Sourav Chakraborty, Sumeer Goyal, D.Roy Choudhury, Unconstrained handwritten character recognition based on fuzzy logic, *Pattern Recognition*, Volume 36, Issue 3, 2003, Pages 603-623, ISSN 0031-3203.
- [10] P. M. Patil, P. S. Dhabe, U. V. Kulkarni and T. R. Sontakke, "Recognition of handwritten characters using modified fuzzy hyperline segment neural network," *The 12th IEEE International Conference on Fuzzy Systems, 2003. FUZZ '03.*, St. Louis, MO, USA, 2003, pp. 1418-1422 vol.2.
- [11] A. V. Nandedkar, K. Venishetti and A. K. Rathod, "Fuzzy min-max neural network based translation, rotation and scale invariant character recognition using RTSI features," *The Fourth International Conference on Computer and Information Technology, 2004. CIT '04.*, Wuhan, China, 2004, pp. 159-164.
- [12] P. M. Patil and M. P. Deshmukh, "Modular general fuzzy hyperline segment neural network," *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, Montreal, QC, Canada, 2005, pp. 1912-1917 vol. 3.