



An Effective Approach for Noise Robust and Rotation Invariant Handwritten Character Recognition Using Zernike Moments Features and Optimal Similarity Measure

Chandan Singh & Ashutosh Aggarwal

To cite this article: Chandan Singh & Ashutosh Aggarwal (2020) An Effective Approach for Noise Robust and Rotation Invariant Handwritten Character Recognition Using Zernike Moments Features and Optimal Similarity Measure, Applied Artificial Intelligence, 34:13, 1011-1037, DOI: [10.1080/08839514.2020.1796370](https://doi.org/10.1080/08839514.2020.1796370)

To link to this article: <https://doi.org/10.1080/08839514.2020.1796370>



Published online: 17 Aug 2020.



Submit your article to this journal [↗](#)



Article views: 396



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 2 View citing articles [↗](#)



An Effective Approach for Noise Robust and Rotation Invariant Handwritten Character Recognition Using Zernike Moments Features and Optimal Similarity Measure

Chandan Singh^a and Ashutosh Aggarwal^{a,b}

^aDepartment of Computer Science, Punjabi University, Patiala, India; ^bDepartment of Computer Science and Engineering, Thapar Institute of Engineering and Technology, Patiala

ABSTRACT

Zernike moments (ZMs) are very effective orthogonal rotation invariant moments. Conventionally, the magnitudes of ZMs are used as feature descriptors and the Euclidean distance is used as a classifier. Recently, a few classifiers based on ZM magnitude and phase have been developed which are reported to be very effective in pattern matching problems. One such a recently developed similarity measure, known as optimal similarity measure, is known to provide very good performance over the ZM magnitude-based Euclidean distance measure in pattern recognition problems, especially under noisy conditions. In this paper, we investigate the conventional magnitude-based similarity measure and the new similarity measures including the optimal similarity measure and compare their performance on segmented handwritten characters and numerals. It is observed that the performance of optimal similarity measure is far better than all other similarity measures. Its performance is very much better than other similarity measures even under very high noisy condition. However, it is slow owing to the optimization of the process involved in its computation. Therefore, we also propose a fast algorithm for its computation and reduce its time complexity. Detailed experimental results are provided to support the above observations.

Introduction

Orthogonal rotation invariant moments (ORIMs) are a class of rotation invariant moments (RIMs) whose kernel functions are orthogonal, meaning thereby that given a set of ORIMs a signal can be reconstructed. The orthogonality property is an important characteristic for describing the image features uniquely. This leads to minimum information redundancy and a set of few ORIMs coefficients are sufficient to represent the image uniquely.

There are many ORIMs in the literature, prominent among them are the Zernike moments (*ZMs*) (Teague 1980), pseudo-Zernike moments (*PZMs*) (Bhatia and Wolf 1954), orthogonal Fourier-Mellin moments (*OFMMs*)

(Sheng and Shen 1994), radial harmonic Fourier moments (*RHRMs*) (Ren et al. 2007), and Chebyshev–Fourier moments (*CHFMs*) (Ping, Wu, and Sheng 2002). A family of these moments has recently been derived from generic Fourier moments (Hoang and Tabbone 2013; Ping et al. 2007). There are many useful characteristics of the ORIMs which make them widely used in numerous image processing applications including printed and handwritten optical character recognition (OCR) system (Bailey and Srinath 1996; Broumandnia and Shanbehzadeh 2007; Kan and Srinath 2002; Khotanzad and Hong 1990; Patil and Sontakke 2007; Ramteke and Mehrotra 2006; Trier, Jain, and Taxt 1996). Khotanzad and Hong (Khotanzad and Hong 1990) are perhaps the first to introduce ZM features to character recognition. They observed that the ZM features are very effective not only for the recognition of the printed characters but also for the recognition of handwritten alphanumeric Roman characters. An exhaustive analysis has been performed by Bailey and Srinath (Bailey and Srinath 1996) on the performance of ZMs and PZMs on the handwritten characters using different classifiers. Experimental results conducted on large databases for unconstrained handwritten numerals reveal that the different variability due to writing style, shape, stroke, and orientation can be handled successfully by their proposed features. Kan and Srinath (Kan and Srinath 2002) have conducted extensive experiments on handwritten characters using ZMs and OFMMs as features and observed that OFMMs provide better recognition performance for small character images. Ramteke and Mehrotra (Ramteke and Mehrotra 2006) evaluated the performance of various techniques based on moment invariants on handwritten Devanagari numerals and observed that the performance of the ZM features is far better than other invariants. Patil and Sontakke (Patil and Sontakke 2007) have used ZM features and fuzzy neural network classifiers for the recognition of Devanagari numerals achieving a very high recognition rate.

The above discussion reveals that ORIMs based feature extraction methods are very useful for orientation invariant handwritten character recognition. They provide robust features against geometric transformation, illumination variation, and noisy conditions. They are used successfully to deal with various issues that are peculiar to handwritten text such as writing style, stroke, orientation, and size. They need little effort for image normalization as compared to other techniques because they are computed on a unit disk and provide global characteristics of the character shape.

Among the various ORIMs the performance of ZMs has proven to be the best in many pattern recognition problems. It has been observed recently by Singh and Upneja (Singh and Upneja 2013) that ZMs are more robust than PZMs and OFMMs. The existing approaches based on ZM, PZM, and OFMM features consider the magnitude of moments for the similarity measure because the magnitude of the moment is invariant under rotation. The

phase of the moments is not rotation invariant and hence it is not included in the comparison of features. It is shown by Oppenheim and Lim (Oppenheim and Lim 1981) that the phase information is more important than the magnitude of moments during signal reconstruction. Keeping this in view, recently, many attempts have been made to incorporate phase information in the similarity measure for the comparison of two patterns. By doing so, considerable improvements in the performance of ORIMs have been reported by many authors in various pattern recognition problems. A new similarity measure called optimal similarity measure involving both magnitude and phase has been developed by Revaud et al. (2009). The method is developed for the orthogonal rotation invariant moments. Motivated by the superior performance of the optimal similarity measure, Singh and Aggarwal (Singh and Aggarwal 2014), recently extended the approach to radial moments which are non-orthogonal. It is shown by them that the performance of the optimal similarity measure even for the non-orthogonal radial moments is very high as compared to the other similarity measures, especially under noisy condition. The optimal similarity measure minimizes a distance function which is defined in terms of magnitude and phase angle of ZMs of two patterns. The distance is very small (ideally zero but because of digital nature, it is non-zero) if the two images are similar. If the two images are different, then the distance value is high. A comparison with the classical ZM magnitude-based similarity measure and many other state-of-the-art techniques have been proposed for image retrieval, and sketch-based 2D and 3D object recognition problem. The performance of the new similarity measure has been observed to be far superior for these applications. The optimal similarity measure is particularly robust against geometric deformation and noise. Chen and Sun (Chen and Sun 2010) have developed a ZMs phase-based descriptor for image matching. Instead of using ZMs magnitude as features, they use phase angles of moments as features. Since the phase is not rotation invariant, their approach first estimates rotation angle between two images and then the relationship between phases of two images is used for the comparison purpose. The success of this method depends upon the accuracy of the estimated rotation angle, if any, between two images. If two images are similar but rotated versions of each other, then the phase distance is zero. If the two images are dissimilar then a fictitious rotation angle will be estimated which will lead to a larger distance between two images. The approach developed by Li et al. (Li, Lee, and Pun 2009) uses both magnitude and phase distances of moments. The phase distance is obtained by estimating the rotation angle between two images from the phase relationship. In the approach developed by Singh et al. (Singh, Walia, and Mittal 2011), the real and imaginary components of ZMs of two images are compared. Their approach does not require the estimation of the rotation angle. They use the phase relationship of ZMs of two images to correct the moments of one of the two images before comparing their real and

imaginary components. L_1 -norm is used as a distance measure. Experimental results on face recognition and printed character recognition problems demonstrate the superiority of their proposed method.

In view of the above newly developed ZMs based similarity measures, the purpose of the present work is to analyze their performance on the handwritten character recognition problem. It is shown that the optimal similarity measure developed by Revaud et al. (Revaud, Lavoue, and Baskurt 2009) outperforms all other similarity measures based on ZM magnitude and phase. We observe that the improvement in the performance using their approach is much more significant than the authors have reported in their paper for the comparison of objects, especially under noisy condition. There are two computational frameworks for the derivation of ZMs: the outer unit disk and the inner unit disk. The outer unit disk encloses the complete image when a transformation is performed for converting a square digital image into a unit disk. In the inner unit disk mapping, the circular disk is contained within the image. It has been observed (Singh and Upneja 2013) that the outer unit disk framework provides more accurate, robust, and stable values of moments compared to its inner unit disk counterpart. We further enhance the performance of the optimal similarity measure using outer unit disk mapping for the computation of ZMs. The major problem with the method of (Revaud, Lavoue, and Baskurt 2009) is its high computation time which is needed for the estimation of the rotation angle. The authors have presented methods for its fast computation. We propose a very fast algorithm that reduces the time complexity of the algorithm. This is supported by a time complexity analysis. A detailed performance analysis has been carried out on the recognition of handwritten characters using various ZM-based similarity measures under the two computational frameworks. In addition to the superior performance under rotation, the proposed method provides excellent recognition results under noisy condition.

The rest of the paper is organized as follows. An overview of the ZMs and its computational framework is provided in Section 2. Section 3 presents the various similarity measures based on ZMs. Fast computation of the optimal similarity measure is developed in Section 4. Detailed experimental results performed on three handwritten character databases, MNIST numeral database, Gurumukhi character database, and Gurumukhi numeral database, are discussed in Section 5. Conclusions and future work are presented in Section 6.

The Zernike Moments (ZMs)

The ZMs of a two-dimensional image function $f(x, y)$ with an order p and repetition q are defined on a unit disk as follows.

$$Z_{pq} = \frac{p+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x, y) V_{pq}^*(x, y) dx dy \quad (1)$$

where $(p, q) \in S = \{ (p, q) \mid 0 \leq p \leq \infty, |q| \leq p, p - |q| = \text{even} \}$.

The kernel function $V_{pq}^*(x, y)$ is the complex conjugate of the Zernike function $V_{pq}(x, y)$ which itself is

$$V_{pq}(x, y) = R_{pq}(x, y) e^{jq\theta} \quad (2)$$

where

$$R_{pq}(x, y) = \sum_{s=0}^{(p-|q|)/2} (-1)^s \frac{(p-s)!}{s! \left(\frac{p+|q|}{2} - s\right)! \left(\frac{p-|q|}{2} - s\right)!} r^{p-2s} \quad (3)$$

with $j = \sqrt{-1}$, and $\theta = \tan^{-1}(y/x)$, $\theta \in [0, 2\pi]$. The Zernike functions are orthogonal on a unit disk, that is,

$$\iint_{x^2+y^2 \leq 1} V_{pq}(x, y) V_{\hat{p}\hat{q}}^*(x, y) dx dy = \frac{\pi}{p+1} \delta_{p\hat{p}} \delta_{q\hat{q}} \quad (4)$$

where δ_{ab} is Kronecker delta defined by

$$\delta_{ab} = \begin{cases} 1, & a = b \\ 0, & \text{otherwise} \end{cases}$$

The orthogonality of the kernel functions is a very important property which leads to minimum information redundancy among the moment values. Consequently, given a set of moments up to a given order p_{max} , the image function can be reconstructed as follows.

$$\hat{f}(x, y) = \sum_{p=0}^{p_{max}} \sum_{\substack{q=-p \\ p-|q| = \text{even}}}^p Z_{pq} V_{pq}(x, y) \quad (5)$$

Invariant Properties of ZMs

If an image is rotated by an angle θ_0 , the ZMs of the original image Z_{pq} and rotated image Z_{pq}^r are related by (Khotanzad and Hong 1990)

$$Z_{pq}^r = Z_{pq} e^{-jq\theta_0} \quad (6)$$

The phase angles φ_{pq} and φ_{pq}^r are related by

$$\varphi_{pq}^r = \varphi_{pq} - q\theta_0 \quad (7)$$

It is observed from Equation (6) that $|Z'_{pq}| = |Z_{pq}|$, meaning thereby that the magnitude of moments is rotation invariant. However, the phases are not rotation invariant as shown in Equation (7). Therefore, the classical pattern matching problems based on ZM use the magnitude of moments as features. The scaling invariance is achieved because the ZMs are computed over a unit disk after performing a mapping process which is explained in the following sub-section. The translation invariance is achieved when the center of the unit disk is placed at the center of mass of the image.

ZMs for Digital Images

The ZMs defined by Equation (1) pertain to a continuous signal $f(x, y)$. In digital image processing, the image function is discrete. If the image function is represented by $f(i, k)$, then i represent the row number and k represents the column number of a pixel (i, k) . Let the size of the image be $N \times N$, then the following transformation converts a $N \times N$ digital square domain into a unit disk:

$$x_i = \frac{2i + 1 - N}{D}, \quad y_k = \frac{2k + 1 - N}{D}, \quad i, k = 0, 1, \dots, N - 1 \quad (8)$$

where D is the digital diameter of the disk given by

$$D = \begin{cases} N & \text{for inner unit disk} \\ N\sqrt{2} & \text{for outer unit disk} \end{cases} \quad (9)$$

The coordinate (x_i, y_k) provides the center of the pixel (i, k) which itself occupies the area $[x_i - \frac{\Delta x}{2}, x_i + \frac{\Delta x}{2}] \times [y_k - \frac{\Delta y}{2}, y_k + \frac{\Delta y}{2}]$, where

$$\Delta x = \Delta y = \frac{2}{D} \quad (10)$$

The above mapping process leads to a simple form of ZMs which are based on zeroth-order approximation of the double integration of Equation (1)

$$Z_{pq} = \frac{4(p+1)}{\pi D^2} \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} f(i, k) V_{pq}^*(x_i, y_k) \quad (11)$$

$$x^2 + y^2 \leq 1$$

The condition $x^2 + y^2 \leq 1$ in Equation (11) is redundant if we compute the moments with the outer unit disk. This mapping of image data over unit disk makes the ZMs to be invariant to scale and rotation.

Similarity Measures

ZMs are complex numbers that comprise two values: magnitude and phase. The classical techniques for character recognition using ZMs consider the magnitude of moments as features as they are invariant to rotation. The phase component of the complex moments is not rotation invariant; therefore, it is not considered during the matching process. Similarly, the real and imaginary components of the moments cannot be equated as the magnitude of their individual components changes under image rotation. The loss of information due to neglecting the phase component reduces the performance of the similarity measure in a matching process. It is shown by Shao and Celenk (Shao and Celenk 2001) that the performance of a shape descriptor increases significantly if both phase and magnitude components are included in the feature set. In view of the importance of the phase information, recently several researchers have attempted to develop similarity measures which are either based on magnitude and phase (Li, Lee, and Pun 2009; Revaud, Lavoue, and Baskurt 2009; Singh and Aggarwal 2014; Singh, Walia, and Mittal 2011) or solely on phase (Chen and Sun 2010). An overview of these measures is discussed in this section.

ZMs Magnitude Based Measure: The Classical Approach

The ZM magnitude-based similarity measure is the classical approach which is widely used for matching of two images. Normally, in the classical approach, Euclidean distance between the magnitudes of ZMs of two images is computed as follows:

$$d_e = \left[\sum_{p=0}^{p_{\max}} \sum_{q=0}^p \left(\left| Z_{pq}^T \right| - \left| Z_{pq}^D \right| \right)^2 \right]^{1/2} \quad (12)$$

$$p - |q| = \text{even}$$

where the subscripts T and D are used, respectively, for the test (query) image and the training image in a database. The major drawback of this measure is that it does not include the phase component of the moments. However, it is very simple to use and computationally very fast.

ZMs Phase-Based Measure

Chen and Sun (Chen and Sun 2010) have derived a phase-based descriptor by estimating the rotation angle between two images using a more accurate method as compared to the simple approach adopted by Li et al. (Li, Lee, and Pun 2009). In their approach, the rotation angle θ_0 is obtained by using the adjacent phase relationship:

$$\theta_0 = \left(\varphi_{p,q} - \varphi_{p,q-1} \right) \bmod 2\pi, \quad q \neq 0 \quad (13)$$

Since $q = 1, 2, 3, \dots, p_{\max}$, $p = 1, 2, 3, \dots, p_{\max}$, $p - q = \text{even}$, there are $\sum_{q=1}^{p_{\max}} (\lfloor p_{\max} - q/2 \rfloor + 1)$ ways to compute the rotation angle θ_0 . A more accurate and robust estimation of the rotation angle is to weigh the estimated angles by the magnitude of ZMs. The complete procedure is described in (Chen and Sun 2010). After estimating rotation angle the phase-based distance is obtained as

$$d_\varphi = \frac{1}{\pi} \sum_{p=0}^{p_{\max}} \sum_{q=0}^p w_{pq} \times \min \left(\left| \varphi_{pq}^C - \varphi_{pq}^D \right|, 2\pi - \left| \varphi_{pq}^C - \varphi_{pq}^D \right| \right) \quad (14)$$

where w_{pq} is the ZM-based weights

$$w_{pq} = \frac{\left| Z_{pq}^D \right| + \left| Z_{pq}^T \right|}{\sum_{p=0}^{p_{\max}} \sum_{q=0}^p \left(\left| Z_{pq}^D \right| + \left| Z_{pq}^T \right| \right)} \quad (15)$$

and φ_{pq}^C is the corrected phase of the test image, $\varphi_{pq}^C = \varphi_{pq}^T + q\theta_0$, and φ_{pq}^D represents the phase of the database image. The method provides a significant improvement in image retrieval as reported in (Chen and Sun 2010).

ZM Magnitude- and Phase-Based Measure

The magnitude- and phase-based descriptor is developed by Li et al. (Li, Lee, and Pun 2009). Their distance measure is a weighted sum of Euclidean distances of magnitude and phase. The Euclidean distance for ZM magnitude, denoted by d_m is obtained by

$$d_m = \left[\sum_{p=0}^{p_{\max}} \sum_{q=0}^p \left(\frac{\left| Z_{pq}^T \right| - \left| Z_{pq}^D \right|}{\max \left(\left| Z_{pq}^T \right|, \left| Z_{pq}^D \right| \right)} \right)^2 \right]^{1/2} \quad (16)$$

In order to compute Euclidean distance of phases, the phase relationship given by Equation (7) is used. For this purpose, it is assumed that the two patterns being matched are rotated versions of each other. The rotation angle is estimated by using Equation (7) by fixing $q = 1$ and taking an odd value of $p \geq 3$ as shown below

$$\theta_0 = \varphi_{p,1}^D - \varphi_{p,1}^T, \quad p \in \{1, 3, 5, \dots, p_{\max}\}. \quad (17)$$

Before matching the phases of two images, the phase of the test image is corrected. Let the corrected phase be denoted by φ^C , then

$$\varphi_{pq}^C = \varphi_{pq}^T + q\theta_0, \quad (p, q) \in S, \quad \varphi_{pq}^C \in [0, 2\pi] \quad (18)$$

The corrected phase φ_{pq}^C of the image T and the phase of the image D φ_{pq}^D have the following relationship

$$\varphi_{pq}^C - \varphi_{pq}^D = \begin{cases} 0 & \text{if } T \text{ and } D \text{ are similar} \\ \text{non-zero} & \text{if } T \text{ and } D \text{ are dissimilar} \end{cases} \quad (19)$$

Therefore, the Euclidean distance of the phases of the database image φ_{pq}^D and the corrected phase of the test image φ_{pq}^C is obtained by

$$d_\varphi = \left[\sum_{p=0}^{p_{\max}} \sum_{q=0}^p \left(\frac{\left(\varphi_{pq}^C - \varphi_{pq}^D \right)}{\pi} \right)^2 \right]^{1/2} \quad (20)$$

The weighted Euclidean distance of the magnitude and the phase is given by

$$d_{m\varphi} = (w_1 d_m + w_2 d_\varphi) / (w_1 + w_2) \quad (21)$$

where w_1 and w_2 are weights to be set empirically. Equal weights have been set in their experimental results (Li, Lee, and Pun 2009).

Similarity Measure Based on Real and Imaginary Components of ZMs

ZMs are complex numbers. Two complex numbers are equal if their real and imaginary components are equal. This fact has been used by Singh et al. (Singh, Walia, and Mittal 2011) while deriving L_1 -norm-based similarity measure between two images. Since only the magnitude of ZMs remains invariant under image rotation and the phase undergoes a change, they correct the phase before applying the similarity measure. They use Equation (6) to correct the ZMs of the test images Z_{pq}^T . Let the corrected ZMs of the image T be denoted by Z_{pq}^C , then Z_{pq}^C can be derived as

$$Z_{pq}^C = Z_{pq}^T e^{jq\theta_0} \quad (22)$$

where from Equation (7) $q\theta_0 = \varphi_{pq}^D - \varphi_{pq}^T$. Therefore, the relationships between the ZMs of the database image D and the corrected ZMs of the test image T are given by

$$\begin{aligned} Z_{pq}^D &= Z_{pq}^C & \text{if two images are similar} \\ Z_{pq}^D &\neq Z_{pq}^C & \text{if two images are dissimilar} \end{aligned} \quad (23)$$

The L_1 -norm based distance is

$$d_{L_1} = \sum_p \sum_q \left(\left| \operatorname{Re}(Z_{pq}^D) - \operatorname{Re}(Z_{pq}^C) \right| + \left| \operatorname{Im}(Z_{pq}^D) - \operatorname{Im}(Z_{pq}^C) \right| \right) \quad (24)$$

It is interesting to observe that if L_2 -norm was used in place of L_1 -norm in the above equation, then the equation would have turned out to be the same as the Euclidean distance between their magnitudes. This approach has a distinct advantage over the other techniques involving both the magnitude and phase, because it does not require the estimation of the rotation angle between two images.

Optimal Similarity -Based Measure

The optimal similarity-based measure is a new similarity-based measure developed by Revaud et al. (Revaud, Lavoue, and Baskurt 2009). It assumes that the two images being matched are similar but rotated versions of each other. Let Z_{pq}^D and Z_{pq}^T denote the ZMs of the two images. The Euclidean distance between the images D and T can be expressed as a function of rotation angle θ as follows (Revaud, Lavoue, and Baskurt 2009).

$$d(\theta) = \sum_{p=0}^{p_{\max}} \sum_{q=-p}^p \frac{\pi}{p+1} \left[\left| Z_{pq}^D \right|^2 + \left| Z_{pq}^T \right|^2 - 2 \left| Z_{pq}^D Z_{pq}^T \right| \cos(q\theta + \varphi_{pq}^T - \varphi_{pq}^D) \right] \quad (25)$$

The rotation angle θ is estimated by minimizing the function $d(\theta)$. Let $\theta = \theta_0$ be a solution to the minimization problem. If the two images are similar, then the value of $d(\theta_0)$ is much smaller than its value for the two dissimilar images.

Fast Computation of Optimal Similarity Based Measure

The optimal similarity-based measure provides a significant improvement in performance over the classical ZM magnitude-based similarity measure as demonstrated in (Revaud, Lavoue, and Baskurt 2009). However, it suffers from a high computation cost. The main reason for the high time complexity is due to the process of minimization of the distance function $d(\theta)$ given by Equation (25). For a given p_{\max} , there are p_{\max} local minima and p_{\max} local maxima in the interval $[0, 2\pi]$. The method requires the global minimum which is computation intensive if the conventional approaches are used. Thus, in order to find the global minimum with less time, a simple and effective solution is given in (Revaud, Lavoue, and Baskurt 2009). The authors have outlined a few steps for its fast computation. We have observed that the algorithm still takes considerable time for the solution of the minimization problem. Here, we present a fast method for its computation, and later, in the

experimental section, we shall present the time complexity analysis. For computational efficiency, we rewrite Equation (25)

$$d(\varphi) = C - 4\pi \sum_{q=1}^{p_{\max}} (A_q \cos(q\varphi) - B_q \sin(q\varphi)) \quad (26)$$

where

$$C = \pi \sum_{\substack{p=0 \\ p=\text{even}}}^{p_{\max}} \frac{1}{p+1} \left[|Z_{p0}^D|^2 + |Z_{p0}^T|^2 - 2|Z_{p0}^D Z_{p0}^T| \cos(\varphi_{p0}^T - \varphi_{p0}^D) \right] \\ + 2\pi \sum_{q=1}^{p_{\max}} \sum_{\substack{p=q \\ p-q=\text{even}}}^{p_{\max}} \frac{1}{p+1} \left[|Z_{pq}^D|^2 + |Z_{pq}^T|^2 \right], \quad (27)$$

$$A_q = \sum_{p=q}^{p_{\max}} \frac{1}{p+1} \left[\operatorname{Re}(Z_{pq}^T) \operatorname{Re}(Z_{pq}^D) + \operatorname{Im}(Z_{pq}^T) \operatorname{Im}(Z_{pq}^D) \right] \quad (28)$$

$$B_q = \sum_{p=q}^{p_{\max}} \frac{1}{p+1} \left[\operatorname{Re}(Z_{pq}^D) \operatorname{Im}(Z_{pq}^T) - \operatorname{Re}(Z_{pq}^T) \operatorname{Im}(Z_{pq}^D) \right] \quad (29)$$

The global optimum is obtained by finding all zero of the derivative of the function $d(\varphi)$, that is

$$d'(\varphi) = 4\pi \sum_{q=1}^{p_{\max}} q(A_q \sin(q\varphi) + B_q \cos(q\varphi)) = 0. \quad (30)$$

As discussed earlier, there are at most $2p_{\max}$ solutions to Equation (30), p_{\max} solutions each pertaining to local maxima and minima. The minima or maxima lie in the equally spaced intervals $[x_n, x_{n+1}]$ where $x_n = \frac{n\pi}{2p_{\max}}$, $0 \leq n < 4p_{\max}$. A local minimum or the maximum lies in the interval $[x_n, x_{n+1}]$ if $d'(x_n)d'(x_{n+1}) \leq 0$. Otherwise, the solution does not lie in this interval. We have observed in actual practice that the number of intervals satisfying this condition is much less than the upper limit for the number of zeros which is $2p_{\max}$. We carried out experiments by matching 7000 binary and 5000 gray scale images for this purpose. The average number of roots for $p_{\max} = 12$ is 7, and 8, for the binary and grayscale images, respectively, out of a possible value of 24. A root of the function $f(x)$ in the interval $[x_n, x_{n+1}]$ is obtained by Regula-Falsi iterative method which is given by:

$$x_{root} = x_n + \frac{\pi}{2p_{max}} \times \frac{f'(x_n)}{(f'(x_n) - f'(x_{n+1}))} \quad (31)$$

Only one iteration is sufficient to approximate the value of the root. More refinements in the root lead to high time complexity without much gain in the accuracy of the root. We have tried Newton-Raphson method which is a higher order method for finding the root of an equation, but its accuracy in the present case, is not better than the Regula-Falsi method. It is probably because of the sinusoidal behavior of the function for which the Regula-Falsi method is better suited. Let the total number of roots in the interval $[0, 2\pi]$ be L . Keeping in view the fact that $L \ll 4p_{max}$, the high computation complexity arises because of the computation of the derivative function $d'(\varphi)$ at $x_n = \frac{n\pi}{2p_{max}}$, $0 \leq n < 4p_{max}$. Based on the above discussion, we propose a fast algorithm for the computation of the functions $d(\varphi)$ and $d'(\varphi)$.

The algorithm:

- (1) Compute C from Equation (27). Its time complexity is $O(p_{max}^2)$.
- (2) Compute A_q and B_q , $q = 1, 2, 3, \dots, p_{max}$ from Equations (28) and (29) and save them in two tables, A_q and B_q . The time complexity is $O(p_{max})$.
- (3) Compute $\cos(q_i\varphi_j)$ and $\sin(q_i\varphi_j)$ once and for all experiments and save them in two tables, C_{ij} and S_{ij} , $i = 0, 1, 2, \dots, p_{max}$, $j = 0, 1, 2, \dots, 4p_{max} - 1$ and $\varphi_j = j\pi/2p_{max}$. The trigonometric functions are very expensive to compute using library functions. The following recurrence relation is used to compute them more efficiently for $q > 1$.

$$\cos((q + 1)\varphi) = a \cos(q\varphi) - b \sin(q\varphi), \quad (32)$$

$$\sin((q + 1)\varphi) = a \sin(q\varphi) + b \cos(q\varphi) \quad (33)$$

where $q = 0, 1, \dots, p_{max} - 1$, $a = \cos(\varphi)$ and $b = \sin(\varphi)$. Therefore, the library functions are used only once to compute the trigonometric functions and these functions for $1 \leq q < p_{max}$ are obtained by using one addition/subtraction and 2 multiplications. The space requirement is $8p_{max}(p_{max} + 1)$ words. Since the function $d(\varphi)$ and its derivative $d'(\varphi)$ are evaluated at $4p_{max}$ locations (for derivative function $d'(\varphi)$) and L locations (for function $d(\varphi)$) and $L \ll p_{max}$, the time complexity of the algorithm turns out to be of the order $O(p_{max}^2)$.

Experimental Analysis

The five similarity measures have been implemented in Microsoft's Visual C++ 6.0 under Windows environment on a PC with 2.13 GHz CPU and 3GB RAM. The experiments are conducted on three databases (DBs) comprising handwritten numerals and characters. The three DBs are MNIST numerals in Roman, Gurumukhi characters (GurChar), and Gurumukhi numerals (GurNum). The Gurumukhi characters and numerals are written in Gurumukhi script, a script of Punjabi language which is the world's 14th largest spoken language. The characteristics of all three test DBs are described briefly as follows.

- (1) MNIST DB: MNIST numeral DB is a standard database of handwritten digits created by NIST (National Institute of Standards and Technology) (MNIST database 1998). It consists of 70,000 grayscale images of numerals of size 28×28 , and 7,000 images per class. The total number of classes is 10. We have used a subset of the DB consisting of 10,000 images, 1,000 images per class. This is done to reduce the time required for conducting various experiments.
- (2) GurChar DB: This database consists of 35 classes of major Gurumukhi characters, each class comprises 200 binary images of size 32×32 . The size of the DB is 7,000 images.
- (3) GurNum DB: This DB comprises 10 numerals of Gurumukhi script. There are 1500 binary images of size 36×36 belonging to 10 classes, each class consisting of 150 images.

The above DBs represent different variations normally present in handwritten scripts, such as writing style, skew, shape, orientation, etc. Some of the handwritten numerals and characters are shown in Figure 1a–c for MNIST, GurChar, and GurNum DBs, respectively. In addition to the experiments for the recognition of characters under normal condition (subject test DB), we perform experiments for rotation invariance and robustness to image noise.

During the course of experiments, we observed that the ZM phase-based measure (Chen and Sun 2010) does not perform well on binary images and its recognition rate is very low. Therefore, we do not include this similarity measure in our analysis. Thus, the analysis is performed using the four similarity measures. In addition, a powerful classifier SVM (support vector machine) is also used which uses ZM magnitude features. This classifier is selected to demonstrate the best recognition results using ZM magnitude, although SVM classifier is very slow and it requires retraining when a new image is introduced in the training DB. Also, the RBF kernel of SVM requires optimal values of parameters for its best performance. However, our objective is to demonstrate the recognition capability of the new similarity measures

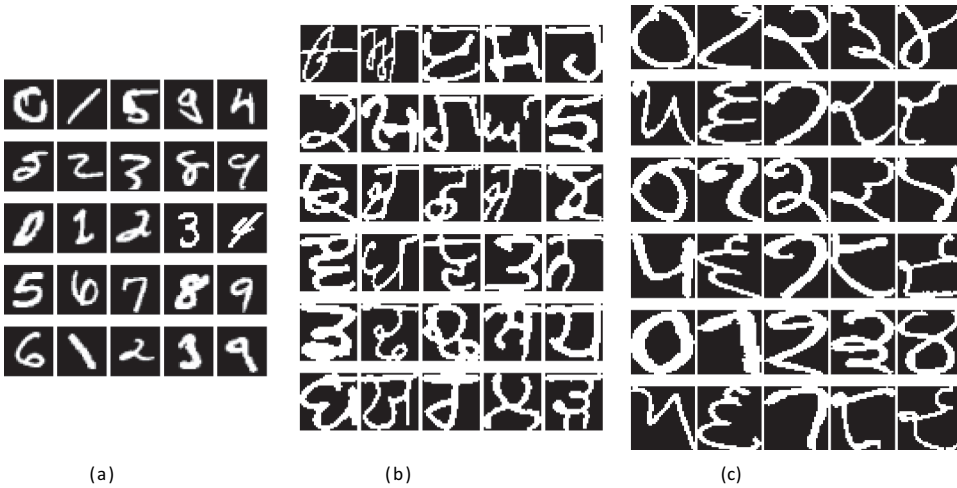


Figure 1. Sample images of three DBs (a) MNIST (Roman numerals), (b) Gurumukhi characters (GurChar), and (c) Gurumukhi numerals (GurNum).

introduced in the recent years and compare their performance with the conventional classifiers. This analysis is very useful to explore the potential of the new similarity measures in various areas where they have not yet been applied. We refer the ZM magnitude-based Euclidean distance measure, ZM magnitude- and phase-based measure, similarity measure based on real and imaginary components of ZMs, optimal similarity-based measure, and ZM magnitude-based measure using SVM classifier as ZM_MAG, ZM_MAG_PHASE, ZM_COMPLEX, ZM_OPT, and ZM_SVM, respectively.

Selection of Number of ZM Features

The low order moments are used to describe an image as they represent its gross aspect and they are more robust to image noise than the high order moments. The recognition performance increases as the number of moment increases. Generally, the increase in the recognition rate becomes more or less stable after a certain moment order. In fact, at high orders, the performance may actually decrease. The time complexity increases quadratically when moment order increases, even if fast recursive methods are used. Also, a large size of feature vector requires more space and high classification time. Therefore, a good tradeoff is needed between the recognition performance and the time and space complexity. For this purpose, we conduct experiments on three subject DBs to examine the moment order versus recognition performance both for the inner unit disk mapping and outer unit disk mapping. The results are shown in [Figure 2a,b](#) for MNIST, in [Figure 3a,b](#) for GurChar, and in [Figure 4a,b](#) for GurNum DBs.

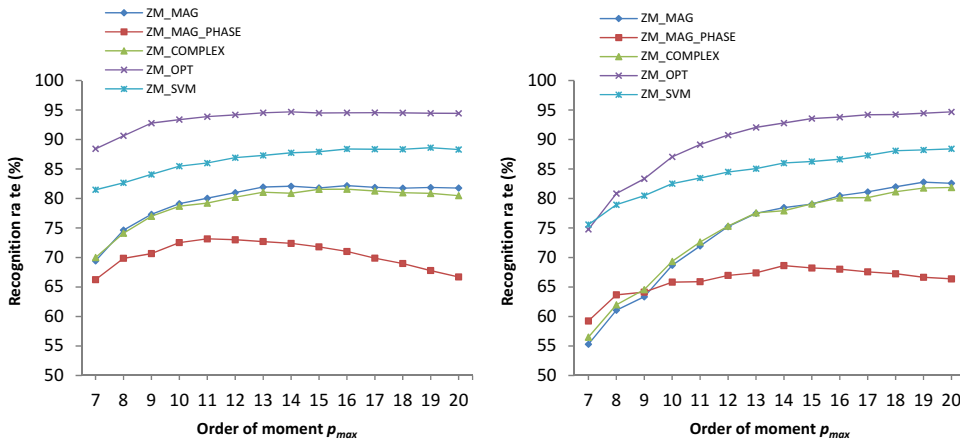


Figure 2. (a) Recognition rate (%) versus moment order for MNIST DB using inner unit disk mapping. (b) Recognition rate (%) versus moment order for MNIST DB using outer unit disk mapping.

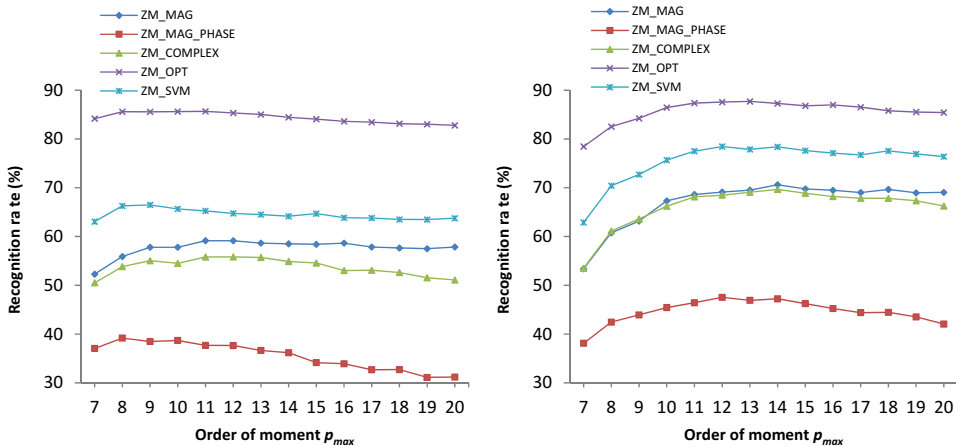


Figure 3. (a) Recognition rate (%) versus moment order for GurChar DB using inner unit disk mapping. (b) Recognition rate (%) versus moment order for GurChar DB using outer unit disk mapping.

We observe from these figures that the recognition rate increases with the increase in moment order. The rate of increase is more at low orders than at high orders and the performance becomes stable after a certain stage. The recognition rate around the moment order 12 is a good compromise keeping in view the space and time complexity. This order is also used by several authors for image matching using ZMs (Chen and Sun 2010; Revaud, Lavoue, and Baskurt 2009; Singh and Aggarwal 2014). Therefore, we use moment order 12 for our experimental analysis. The moment Z_{00} represents the average gray value of an image and Z_{11} is zero if the center of the unit disk is at the center of

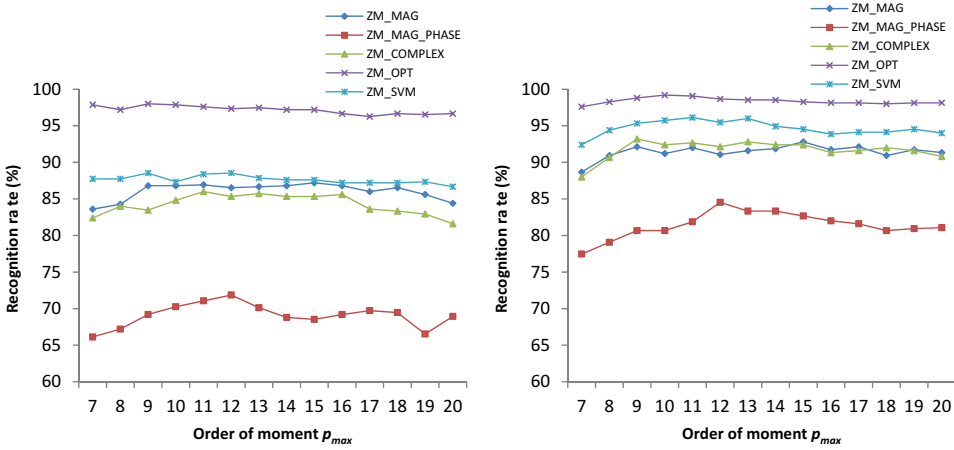


Figure 4. (a) Recognition rate (%) versus moment order for GurNum DB using inner unit disk mapping. (b) Recognition rate (%) versus moment order for GurNum DB using outer unit disk mapping.

the image. Therefore, they are excluded from the set of the feature vector. The total number of moments is 47 for $p_{max} = 12$ after discarding Z_{00} and Z_{11} .

Subject DBs Tests

We divide the three DBs into training and test DBs. The number of images in training and test DBs is equal. Thus, MNIST training and test DBs comprise 5,000 images each. The Gurumukhi character and numeral training and test DBs comprise 3,500 and 750 images, respectively. The selection of images in these DBs is done randomly. The recognition rates for MNIST numerals, Gurumukhi characters, and Gurumukhi numerals are shown in Table 1 for inner unit disk mapping and outer unit disk mapping.

The following results are obtained from these experiments.

- (i) The performance of optimal similarity measure is much better than all other similarity measures. If we compare its performance with ZM magnitude-based Euclidean distance measure, then we note that its performance is higher by 13.18%, 26.20%, and 10.80% for MNIST,

Table 1. Recognition rates (%) for three DBs for inner and outer unit disk mappings.

Method	MNIST		GurChar		GurNum	
	Inner	Outer	Inner	Outer	Inner	Outer
ZM_MAG	81.00	75.24	59.14	69.11	86.53	91.07
ZM_MAG_PHASE (Li, Lee, and Pun 2009)	73.00	66.96	37.66	47.54	71.87	84.53
ZM_COMPLEX (Singh, Walia, and Mittal 2011)	80.22	75.32	55.83	68.49	85.33	92.13
ZM_SVM	86.92	84.50	64.74	78.46	88.53	95.47
ZM_OPT (Revaud, Lavoue, and Baskurt 2009)	94.18	90.76	85.34	87.57	97.33	98.67

GurChar, and GurNum DBs, respectively, for inner unit disk mapping. For the outer unit disk mapping, these improvements are 15.52%, 18.46%, and 7.60%, respectively. The second best recognition rates are obtained by SVM classifier. When compared to SVM classifier, the improvements in the recognition rates using optimal similarity measure for the three DBs are 7.26%, 20.60%, and 8.80%, for the inner unit disk mapping and 6.26%, 9.11%, and 3.20%, for the outer unit disk mapping. The magnitude- and phase-based descriptors, ZM_MAG_PHASE and ZM_COMPLEX, provide recognition rates which are comparable to ZM_MAG similarity measure. One of the reasons for their low performance compared to ZM_OPT is that they do not provide high recognition rates for binary images, although their performance on the recognition of grayscale images is very significant. This result is based on our observation that when rotation angle is estimated between two images by these methods, the estimated rotation angle is very close to the actual rotation angle when the images are gray. On the other hand, when the rotation angle is estimated on the binary images, then the accuracy drops significantly. The optimal similarity-based measure finds rotation angle after solving the global optimization problem, and its performance remains more or less the same irrespective of the fact whether the two matching images are binary or grayscale.

- (ii) The improvement in the recognition rate on Gurumukhi character DB, GurChar, using ZM_OPT is very high. As compared to ZM_MAG, it provides 26.20% and 18.46% more recognition rates for inner unit disk mapping and outer unit disk mapping, respectively. This DB represents wide variations within a class; therefore, the improvements achieved by ZM_OPT are very significant. This reflects the robustness of the optimal similarity measure in classifying objects with large variations within a class.
- (iii) The outer unit disk mapping provides much higher recognition rates compared to the inner unit disk mapping for all DBs except for the MNIST DB. The reason for the low improvement in MNIST DB is attributed to the fact that the actual character shapes are much inside the image boundaries and the space between the character boundaries and image boundaries is filled with the background color. This is clear from the MNIST sample character images shown in [Figure 1a](#). There is not much loss of shape information when an inner unit disk is used. On the other hand, when an outer unit disk is used, irrelevant information is introduced between the character boundaries and image boundaries. This affects the recognition rate. The percentage drop in recognition rate for MNIST DB using ZM_OPT (3.42%) is less than that obtained by ZM_MAG (5.76%), ZM_MAG_PHASE (6.04%), and

ZM_COMPLEX (4.90%), but slightly higher than that obtained by ZM_SVM (2.42%). Since no normalization including the detection of character bounding box is performed in our experiments, these results prove the robustness of the ZM_OPT similarity measure against shape normalization process. In fact the performance of ZM_MAG_PHASE is worst in the outer unit disk mapping. It is due to the fact that the estimation of rotation angle is affected when the character image contains irrelevant information outside the bounding box of the character image which impacts its accuracy.

Rotation DBs Tests

We rotate all test images (5,000 for MNIST, 3,500 for GurChar, and 750 for GurNum) from 10° to 90° with an interval of 10° and also include the rotation at 45° . Rotation at 45° is included because of the 8-way symmetry/anti-symmetry of the kernel function of ZMs (Singh and Walia 2011). The trend in recognition rates is expected to be symmetrical about 45° , but due to digitization error involved during image rotation, this property is affected. The training DBs consist of non-rotated original images as used in the subject DB tests (5,000 for MNIST, 3,500 for GurChar, and 750 for GurNum). The recognition rates are depicted in Tables 2a and 2b for MNIST, in Tables 3a

Table 2a. Recognition rates (%) for rotated Roman numerals (MNIST DB) using inner unit disk mapping.

Method	Angle of rotation										
	0°	10°	20°	30°	40°	45°	50°	60°	70°	80°	90°
ZM_MAG	81.00	80.34	80.08	80.04	79.84	79.56	80.28	80.32	79.96	79.70	81.00
ZM_MAG_PHASE (Li, Lee, and Pun 2009)	73.00	72.12	72.9	73.04	73.14	73.36	72.58	72.76	72.68	72.52	73.00
ZM_COMPLEX (Singh, Walia, and Mittal 2011)	80.22	79.72	79.1	79.80	79.32	79.18	79.78	79.90	79.26	78.96	80.22
ZM_SVM	86.92	86.52	86.38	86.34	86.52	85.82	86.48	86.40	86.34	86.06	86.92
ZM_OPT (Revaud, Lavoue, and Baskurt 2009)	94.18	93.96	93.98	94.20	93.92	93.80	93.84	93.88	94.02	94.00	94.18

Table 2b. Recognition rates (%) for rotated Roman numerals (MNIST DB) using outer unit disk mapping.

Method	Angle of rotation										
	0°	10°	20°	30°	40°	45°	50°	60°	70°	80°	90°
ZM_MAG	75.24	74.44	74.54	73.94	74.34	73.62	73.86	74.40	74.52	74.44	75.24
ZM_MAG_PHASE (Li, Lee, and Pun 2009)	66.96	66.60	66.98	67.28	66.30	66.54	66.14	66.68	66.14	66.92	66.96
ZM_COMPLEX (Singh, Walia, and Mittal 2011)	75.32	74.40	74.78	73.84	74.62	73.26	74.02	74.80	74.44	74.34	75.32
ZM_SVM	84.50	84.04	84.02	84.00	83.98	82.68	83.68	83.86	84.20	83.84	84.50
ZM_OPT (Revaud, Lavoue, and Baskurt 2009)	90.76	89.98	90.46	89.98	90.40	90.14	89.78	89.90	90.58	90.50	90.76

Table 3a. Recognition rates (%) for rotated Gurumukhi characters (GurChar DB) using inner unit disk mapping.

Method	Angle of rotation										
	0°	10°	20°	30°	40°	45°	50°	60°	70°	80°	90°
ZM_MAG	59.14	57.49	57.69	58.20	57.14	57.51	58.03	57.34	57.34	57.46	59.14
ZM_MAG_PHASE (Li, Lee, and Pun 2009)	37.66	35.94	37.37	36.91	36.60	37.26	36.60	38.03	36.94	37.06	37.66
ZM_COMPLEX (Singh, Walia, and Mittal 2011)	55.83	54.31	54.09	53.69	53.66	54.20	53.43	53.86	54.03	53.86	55.83
ZM_SVM	64.74	64.06	63.49	63.11	63.83	63.91	64.09	63.97	63.69	63.71	64.74
ZM_OPT (Revaud, Lavoue, and Baskurt 2009)	85.34	85.06	85.20	85.60	85.09	85.31	85.00	85.20	85.11	85.17	85.34

Table 3b. Recognition rates (%) for rotated Gurumukhi characters (GurChar DB) using outer unit disk mapping.

Method	Angle of rotation										
	0°	10°	20°	30°	40°	45°	50°	60°	70°	80°	90°
ZM_MAG	69.11	68.94	69.37	68.71	68.97	68.83	68.77	69.06	69.63	68.94	69.11
ZM_MAG_PHASE (Li, Lee, and Pun 2009)	47.54	47.20	47.49	46.77	46.14	47.03	46.74	47.80	47.09	48.43	47.54
ZM_COMPLEX (Singh, Walia, and Mittal 2011)	68.49	69.09	68.49	68.26	67.94	68.40	67.89	68.60	68.31	68.71	68.49
ZM_SVM	78.46	77.80	77.71	77.66	77.40	77.14	77.31	77.63	77.83	77.26	78.46
ZM_OPT (Revaud, Lavoue, and Baskurt 2009)	87.57	87.23	87.14	87.40	87.43	86.91	87.09	87.20	87.46	87.37	87.57

Table 4a. Recognition rates (%) for rotated Gurumukhi numerals (GurNum DB) using inner unit disk mapping.

Method	Angle of rotation										
	0°	10°	20°	30°	40°	45°	50°	60°	70°	80°	90°
ZM_MAG	86.53	87.20	86.93	87.47	86.13	86.40	86.80	87.07	86.67	86.93	86.53
ZM_MAG_PHASE (Li, Lee, and Pun 2009)	71.87	69.07	69.60	71.47	71.20	71.20	71.47	69.20	70.53	71.07	71.87
ZM_COMPLEX (Singh, Walia, and Mittal 2011)	85.33	83.87	84.67	84.53	83.73	83.33	83.60	84.27	83.47	84.00	85.33
ZM_SVM	88.53	87.20	87.60	87.20	87.60	87.87	88.13	87.07	86.93	87.07	88.53
ZM_OPT (Revaud, Lavoue, and Baskurt 2009)	97.33	97.20	97.33	97.33	97.07	97.47	97.07	97.47	97.07	97.20	97.33

and 3b for GurChar, and in Tables 4a and 4b for GurNum DBs for inner and outer unit disk mappings.

It is observed from these tables that the recognition rates achieved by optimal similarity measure ZM_OPT are much more than the other four similarity measures. The outer unit disk mapping provides better rotation invariance results than the inner unit disk mapping for GurChar and GurNum DBs while its performance for MNIST is poor, the reason for which has already been explained under Section Subject DBs Tests. It is also observed from the tables that sometimes in a few cases the recognition rate for rotated images is better than the non-rotated images. For instance, a recognition rate of 94.20% at 30° for MNIST DB is better than the recognition rate of 94.18% at

Table 4b. Recognition rates (%) for rotated Gurumukhi numerals (GurNum DB) using outer unit disk mapping.

Method	Angle of rotation										
	0°	10°	20°	30°	40°	45°	50°	60°	70°	80°	90°
ZM_MAG	91.07	90.80	91.20	90.80	90.27	91.07	90.27	91.33	91.47	90.00	91.07
ZM_MAG_PHASE (Li, Lee, and Pun 2009)	84.53	83.20	84.93	82.13	84.53	81.47	80.93	84.27	83.73	83.33	84.53
ZM_COMPLEX (Singh, Walia, and Mittal 2011)	92.13	92.00	92.27	91.60	92.13	92.13	92.53	90.93	91.60	92.00	92.13
ZM_SVM	95.47	95.07	94.80	94.67	94.67	94.93	95.07	95.87	95.07	95.07	95.47
ZM_OPT (Revaud, Lavoue, and Baskurt 2009)	98.67	98.53	98.80	98.80	98.93	98.67	98.40	98.67	98.53	98.53	98.67

0° (non-rotated) which is evident from Table 2a. The actual number of correctly recognized images at 30° is 4710, whereas at 0° rotation it is 4709, out of a total of 5000 test images. The reason for this perceived abnormality is a better alignment between the test images and training images, because some of the non-rotated images in the training DBs appear in some degree of orientation and skew. The ZM_OPT similarity measure provides very high recognition rates on GurChar and GurNum DBs and it is least affected by image rotation. It is observed that the maximum and the minimum recognition rates on GurChar DB for ZM_MAG, ZM_MAG_PHASE, ZM_COMPLEX, ZM_SVM, and ZM_OPT for inner unit disk mapping are (57.14%, 59.14%), (35.94%, 38.03%), (53.43%, 55.83%), (63.11%, 64.74%), and (85.00%, and 85.60%), respectively. The difference between the maximum and the minimum recognition rates under rotation for ZM_OPT is insignificant which proves its robustness against rotation. The difference is only 0.60% as compared to the difference caused by the other similarity measures which ranges from 1.5% to 2.5% approximately. The trend in results is similar in the outer unit disk mapping. The trend in recognition results on GurNum DB is similar to those obtained on the GurChar DB. The recognition rates for GurNum are higher because the number of classes in GurNum is only 10 in comparison to 35 classes in the GurChar DB.

Noisy DBs Tests

All test images (5,000 for MNIST, 3,500 for GurChar, and 750 for GurNum) are subjected to noise test by adding salt-and-pepper noise with noise densities 5% to 25% at an increment of 5%. A few noisy images at these levels are shown in Figure 5a–c, for the three DBs. The recognition rates are depicted in Tables 5a, 5b, and 5c. It is observed that the effect of noise on recognition rates is insignificant using the optimal similarity measure ZM_OPT.

It is observed from Table 5a that the drop in the recognition rates of ZM_OPT on MNIST DB in the case of inner unit disk mapping is 0.48%, 1.46%, 2.16%, 3.40%, and 6.44%, for noise densities 5% through 25% which are

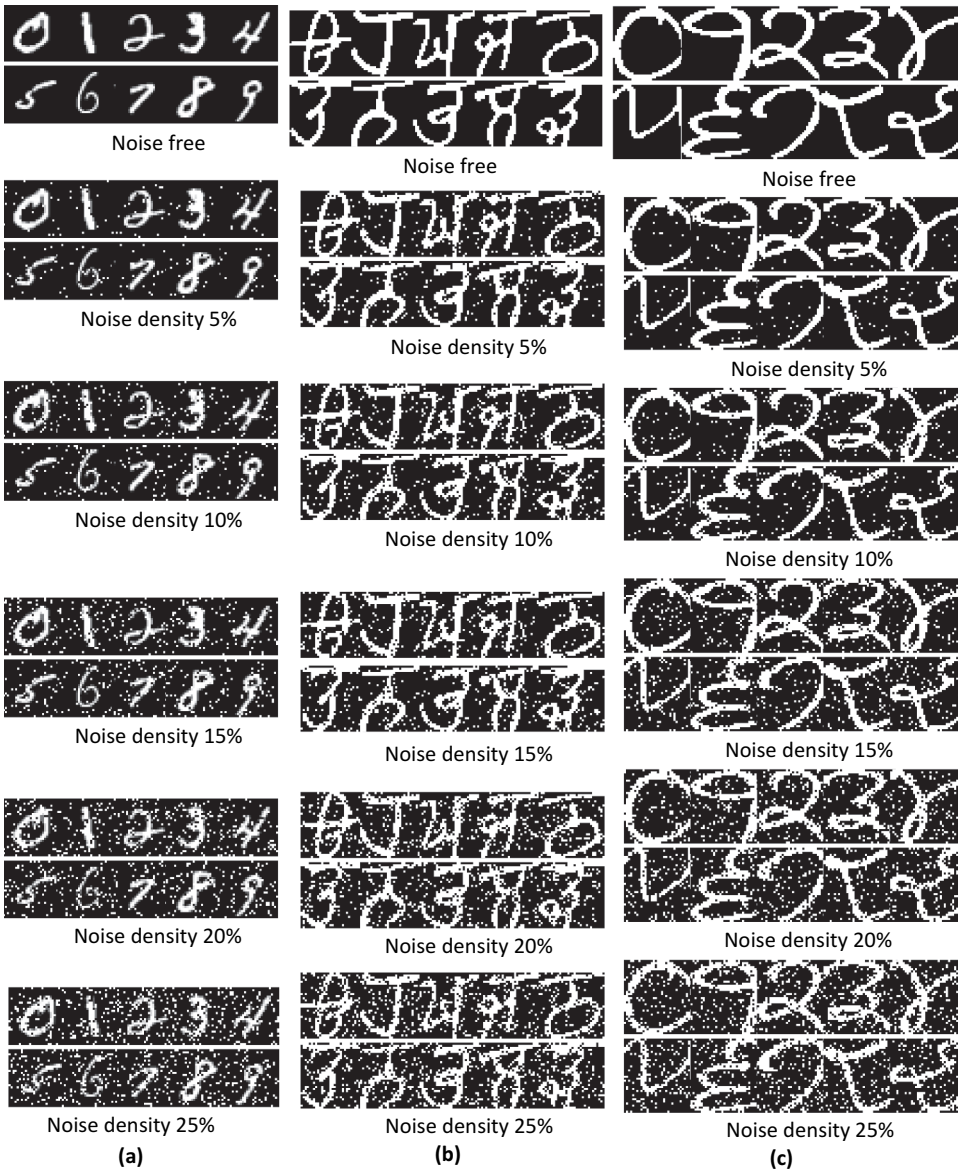


Figure 5. Noise samples of three DBs at different noise densities. (a) MNIST Roman numerals, (b) Gurumukhi characters, and (c) Gurumukhi numerals.

very small. The performance of other similarity measures is very poor for these noise densities. Using inner disk mapping, the drops in recognition rates for ZM_MAG are 3.50%, 7.48%, 14.30%, 21.46%, and 29.48%. For ZM_MAG_PHASE, ZM_COMPLEX, and ZM_SVM these drops are (7.78%, 14.06%, 21.94%, 27.76%, and 33.02%), (3.68%, 8.98%, 15.76%, 25.18%, and 32.24%), and (3.82%, 10.94%, 21.18%, 27.76%, and 33.36%). A similar trend is observed for outer unit disk mapping for all similarity measures, although the drop in recognition rate for outer unit disk mapping for MNIST DB is

Table 5a. Recognition rates (%) for MNIST Roman numerals DB at different noise densities.

Method	Inner unit disk						Outer unit disk					
	Noise-free (0%)	Noise density					Noise-free (0%)	Noise density				
		5%	10%	15%	20%	25%		5%	10%	15%	20%	25%
ZM_MAG	81.00	77.50	73.52	66.70	59.54	51.52	75.24	72.20	65.08	56.02	43.46	35.42
ZM_MAG_PHASE (Li, Lee, and Pun 2009)	73.00	65.22	58.94	51.06	45.24	39.98	66.96	47.14	38.94	32.38	28.78	24.78
ZM_COMPLEX (Singh, Walia, and Mittal 2011)	80.22	76.54	71.24	64.46	55.04	47.98	75.32	70.96	64.56	57.50	49.00	39.78
ZM_SVM	86.92	83.10	75.98	65.74	59.16	53.56	84.50	76.20	63.88	56.18	49.54	42.40
ZM_OPT (Revaud, Lavoue, and Baskurt 2009)	94.18	93.70	92.72	92.02	90.78	87.74	90.76	89.34	87.84	82.60	74.70	66.42

Table 5b. Recognition rates (%) for Gurumukhi characters DB at different noise densities.

Method	Inner unit disk						Outer unit disk					
	Noise-free (0%)	Noise density					Noise-free (0%)	Noise density				
		5%	10%	15%	20%	25%		5%	10%	15%	20%	25%
ZM_MAG	59.14	57.06	54.40	50.80	44.23	37.46	69.11	68.00	64.69	61.26	54.34	46.14
ZM_MAG_PHASE (Li, Lee, and Pun 2009)	37.66	36.74	35.26	33.03	32.00	28.14	47.54	45.71	44.91	43.83	40.34	38.40
ZM_COMPLEX (Singh, Walia, and Mittal 2011)	55.83	53.23	50.74	44.37	38.11	30.09	68.49	66.60	63.34	59.74	52.89	45.09
ZM_SVM	64.74	62.86	60.66	56.60	50.69	43.94	78.46	75.74	72.29	66.40	59.00	49.37
ZM_OPT (Revaud, Lavoue, and Baskurt 2009)	85.34	85.17	84.51	84.00	83.66	83.69	87.57	86.66	86.49	85.43	84.17	82.34

Table 5c. Recognition rates (%) for Gurumukhi numerals DB at different noise densities.

Method	Inner unit disk						Outer unit disk					
	Noise-free (0%)	Noise density					Noise-free (0%)	Noise density				
		5%	10%	15%	20%	25%		5%	10%	15%	20%	25%
ZM_MAG	86.53	85.60	84.13	82.53	79.20	74.93	91.07	91.07	91.07	88.93	87.07	83.60
ZM_MAG_PHASE (Li, Lee, and Pun 2009)	71.87	69.60	69.07	65.33	62.27	59.33	84.53	81.87	78.67	78.40	75.47	75.33
ZM_COMPLEX (Singh, Walia, and Mittal 2011)	85.33	83.60	81.20	80.13	73.20	69.07	92.13	91.73	90.27	89.87	86.53	82.00
ZM_SVM	88.53	86.53	86.00	84.93	81.33	77.47	95.47	95.20	93.20	92.40	89.60	84.80
ZM_OPT (Revaud, Lavoue, and Baskurt 2009)	97.33	97.20	97.20	97.06	96.93	96.93	98.67	98.40	98.40	98.40	98.40	98.13

significant as compared to the inner unit disk mapping. This is attributed to the reason as explained under Section Subject DBs Tests. When we use outer unit disk mapping, the optimal similarity measure ZM_OPT provides lower recognition rates as compared to inner unit disk mapping for MNIST DB. In

fact, the drop in the recognition rate is significant as well. Whereas the drop in recognition rate due to noise for inner unit disk mapping for MNIST DB is from 94.18% to 87.74%, the drop for the outer unit disk mapping is from 90.76% to 66.42%. The drop in recognition rates for GurChar using the two mappings and ZM_OPT similarity measure do not vary much. This shows the high robustness of the ZM_OPT method against image noise.

Time Complexity Analysis

All similarity measures require the computation of all ZMs up to the maximum order p_{max} . Fast algorithms are used to compute ZMs as discussed by Singh and Walia (Singh and Walia 2011). The time complexity of the fast methods is $O(N^2 p_{max}^2)$. The ZMs of the training images are computed offline and saved in the database. Hence, the time taken for testing is crucial as generally the testing is performed online. The total time taken during training and testing by the different similarity measures for $p_{max} = 12$ is shown in Table 6. The average time taken for feature extraction and testing for one image is shown in brackets. It is observed that the time taken during the training is very small as compared to testing. During testing, the time taken for feature extraction for one image is the same as during training.

The average time taken for feature extraction for an image of MNIST, GurChar, and GurNum DBs using inner unit disk mapping is 0.00022 sec, 0.00028 sec, and 0.00037 sec, respectively. The average time taken is different for the three DBs because the image sizes are different. During testing, the time taken for feature extraction is very less as compared to classification. In fact, the major part of the total time taken is due to the classification process. It is observed that the ZM_MAG classifier takes much less time than other similarity measures. The similarity measures ZM_SVM and ZM_OPT are the slowest ones. For small DBs (GurChar and GurNum), the ZM_OPT classifier takes less time as compared to ZM_SVM. For large DB (MNIST) the trend is reversed. For GurChar DB, the time taken by ZM_SVM for inner and outer unit disk mappings is 214.36 sec and 191.02 sec, while for ZM_OPT these values are 166.17 sec and 160.44 sec. For GurNum DB, these values are 17.26 sec, 15.35 sec for ZM_SVM, and 8.10 sec, 7.78 sec for ZM_OPT classifiers. During experiments, it is observed that the classifiers ZM_OPT and ZM_SVM take more time for classification in the inner unit disk mapping than the outer unit disk mapping. For the ZM_OPT classifier, after investigation, it was found that the optimization process using inner unit disk mapping is slower than the outer unit disk mapping. When we solve Equation (30) for finding its zeros, it provided 159115233, 116995884, and 5574910 zeros for the inner unit disk mapping as compared to 127235895, 97915984, and 4396884 zeros for the outer unit disk mapping for the three DBs. It is made clear here that a test image is compared with all the training images and the number of

Table 6. CPU elapse time (in sec) for training and testing on three DBs for $p_{max} = 12$. The quantities in parenthesis represent the average time taken for one image.

Data sets	Testing (feature extraction & classification)												
	Training (feature extraction)					Testing (feature extraction & classification)							
	Inner unit disk		Outer unit disk		ZM_MAG	ZM_MAG_PHASE	ZM_COMPLEX	ZM_SVM	ZM_OPT	ZM_MAG	ZM_MAG_PHASE	ZM_COMPLEX	ZM_SVM
MINIST	1.09 (0.00022)	1.27 (0.00025)	17.88 (0.0036)	129.37 (0.026)	129.37 (0.026)	265.04 (0.053)	244.31 (0.049)	319.06 (0.064)	18.19 (0.004)	127.53 (0.026)	264.03 (0.053)	216.36 (0.043)	311.88 (0.062)
GUR CHAR	0.97 (0.00028)	1.12 (0.00032)	8.22 (0.0024)	64.55 (0.018)	64.55 (0.018)	131.03 (0.037)	214.36 (0.061)	166.17 (0.047)	8.37 (0.002)	64.25 (0.018)	130.78 (0.037)	191.02 (0.055)	160.44 (0.046)
GUR NUM	0.28 (0.00037)	0.31 (0.00041)	0.89 (0.0012)	3.78 (0.005)	3.78 (0.005)	6.41 (0.009)	17.26 (0.023)	8.10 (0.011)	0.95 (0.001)	3.82 (0.005)	6.44 (0.009)	15.35 (0.020)	7.78 (0.010)

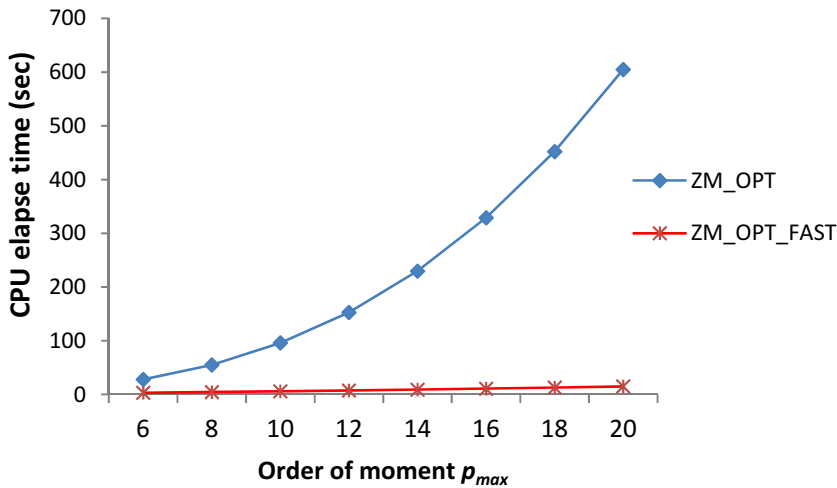


Figure 6. Comparison of CPU elapse time taken by a naïve implementation of ZM_OPT and its proposed fast method.

zeros mentioned here accounts for all test images. The average number of zeros is 6.365, 9.55, and 9.910 for inner unit disk mapping, and 5.089, 7.99, and 7.817, for the outer unit disk mapping for the three DBs. For ZM_SVM, we optimize the parameters for the radial kernel basis function of SVM. This optimization process requires more number of iterations in inner unit disk mapping than in the outer unit disk mapping. Therefore, the outer unit disk mapping takes less computation time than the inner unit disk mapping.

It is clear from the performance and time complexity analysis that the optimal similarity measure ZM_OPT outperforms all other similarity measures and its performance is much better than the next best classifier ZM_SVM. It is much more robust against rotation and noise. Even at a very high noise density, its performance is excellent. However, it suffers from high computation time. The proposed method for its fast computation reduces its time complexity and brings it at par with the existing best classifier ZM_SVM. Figure 6 shows a comparative time performance analysis between a naïve implementation of ZM_OPT and its proposed fast method as a function of p_{max} for the GurNum DB.

Conclusion

It is observed from the experimental analysis that the optimal similarity measure which uses ZM magnitude and phase for finding the minimum distance between two images is an excellent similarity measure whose recognition performance is much better than all other similarity measures. It is very robust to image rotation and noise. Even at a very high noise density of 25% of salt-and-pepper noise, its performance is similar to the noise-free image. The

high recognition rates on three segmented handwritten character databases reflect its robustness against writing style, skew, shape, orientation, etc. It is, however, computation intensive. The proposed fast algorithm reduces its time complexity and brings its time requirement close to the next best classifier which is based on the SVM. Our future work will focus on its use to other applications and further reduction in its time complexity.

Acknowledgments

One of the authors (Ashutosh Aggarwal) is thankful to the University Grants Commission (UGC), Govt. of India, for providing fellowship under Special Assistance Programme of UGC vide file no: F.4-16/2015/DRS-III (SAP-II).

References

- Bailey, R. R., and M. Srinath. 1996. Orthogonal moment features for use with parametric and non parametric classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (4):389–99. doi:10.1109/34.491620.
- Bhatia, A. B., and E. Wolf. 1954. On the circle polynomials of Zernike and related orthogonal sets. *Mathematical Proceedings of the Cambridge Philosophical Society* 50 (1):40–48. doi:10.1017/S0305004100029066.
- Broumandnia, A., and J. Shanbehzadeh. 2007. Fast Zernike wavelet moments for Farsi character recognition. *Image and Vision Computing* 25 (5):717–26. doi:10.1016/j.imavis.2006.05.014.
- Chen, Z., and S.-K. Sun. 2010. A Zernike moment phase-based descriptor for local image representation and matching. *IEEE Transactions on Image Processing* 19 (1):205–19. doi:10.1109/TIP.2009.2032890.
- Hoang, T. V., and S. Tabbone. 2013. Errata and comments on “Generic orthogonal moments: Jacobi–Fourier moments for invariant image description”. *Pattern Recognition* 46 (11):3148–55. doi:10.1016/j.patcog.2013.04.011.
- Kan, C., and M. D. Srinath. 2002. Invariant character recognition with Zernike and orthogonal Fourier-Mellin moments. *Pattern Recognition* 35 (1):143–54. doi:10.1016/S0031-3203(00)00179-5.
- Khotanzad, A., and Y. H. Hong. 1990. Invariant Image recognition by Zernike moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12 (5):489–97. doi:10.1109/34.55109.
- Li, S., M.-C. Lee, and C. M. Pun. 2009. Complex Zernike moments features for shape based image retrieval. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* 39 (1):227–37. doi:10.1109/TSMCA.2008.2007988.
- MNIST database. 1998. <http://yann.lecun.com/exdb/mnist>.
- Oppenheim, A. V., and J. S. Lim. 1981. The importance of phase in signals. *Proceedings of the IEEE* 69 (5):529–41. doi:10.1109/PROC.1981.12022.
- Patil, P. M., and T. R. Sontakke. 2007. Rotation, scale and translation invariant Devanagari numeral character recognition using general fuzzy neural network. *Pattern Recognition* 40 (7):2110–17. doi:10.1016/j.patcog.2006.12.018.

- Ping, Z., H. Ren, J. Zou, Y. Sheng, and W. Bo. 2007. Generic orthogonal moments: Jacobi–Fourier moments for invariant image description. *Pattern Recognition* 40 (4):1245–54. doi:10.1016/j.patcog.2006.07.016.
- Ping, Z. L., R. G. Wu, and Y. L. Sheng. 2002. Image description with Chebyshev-Fourier moments. *Journal of the Optical Society of America* 19 (9):1748–54. doi:10.1364/JOSAA.19.001748.
- Ramteke, R. J., and S. C. Mehrotra. 2006. Feature Extraction based on moment invariants for handwriting recognition. IEEE Conference on Cybernetics and Intelligent Systems, Bangkok, Thailand, 1–6.
- Ren, H., A. Liu, J. Zou, D. Bai, and Z. Ping. 2007. Character reconstruction with radial harmonic Fourier moments. 4th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD07), Haikou, China, 307–10.
- Revaud, J., G. Lavoue, and A. Baskurt. 2009. Improving Zernike moments comparison for optimal similarity and rotation angle retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31 (4):627–36. doi:10.1109/TPAMI.2008.115.
- Shao, Y., and M. Celenk. 2001. Higher order spectra (HOS) invariants for shape recognition. *Pattern Recognition* 34 (11):2097–113. doi:10.1016/S0031-3203(00)00148-5.
- Sheng, Y., and L. Shen. 1994. Orthogonal Fourier-Mellin moments for invariant pattern recognition. *Journal of the Optical Society of America* 11 (6):1748–57. doi:10.1364/JOSAA.11.001748.
- Singh, C., and A. Aggarwal. 2014. A noise resistant image matching method using angular radial transform. *Digital Signal Processing* 33:116–24. doi:10.1016/j.dsp.2014.07.004.
- Singh, C., and R. Upneja. 2013. Error analysis in computation of orthogonal rotation invariant moments. *Journal of Mathematical Imaging and Vision* 1–21. doi:10.1007/s10851-013-0456-1.
- Singh, C., and E. Walia. 2011. Algorithms for fast computation of Zernike moments and their numerical stability. *Image and Vision Computing* 29 (4):251–59. doi:10.1016/j.imavis.2010.10.003.
- Singh, C., E. Walia, and N. Mittal. 2011. Rotation invariant complex Zernike moments features and their applications to human face and character recognition. *IET Computer Vision* 5 (5):255–66. doi:10.1049/iet-cvi.2010.0020.
- Teague, M. R. 1980. Image analysis via the general theory of moments. *Journal of the Optical Society of America* 70 (8):920–30. doi:10.1364/JOSA.70.000920.
- Trier, O. D., A. K. Jain, and T. Taxt. 1996. Feature extraction methods for character recognition — a survey. *Pattern Recognition* 29 (4):641–62. doi:10.1016/0031-3203(95)00118-2.