Robust Moment Invariant with Higher Discriminant Factor Based on Fisher Discriminant Analysis for Symbol Recognition

Widya Andyardja Weliamto[1], Hock Soon Seah[1], and Antonius Wibowo[2]

[1] School of Computer Engineering, Nanyang Technological University,
Block N4, Nanyang Avenue, Singapore 639798
{ph810023, ashsseah}@ntu.edu.sg
http://www.ntu.edu.sg/sce
[2] Department of Electrical Engineering, Bandung Institute of Technology,
Ganesha 10, Indonesia 40132
antonius232@students.itb.ac.id

Abstract. In this paper, we propose a robust moment invariant which has a higher discriminant factor based on Fisher linear discriminant analysis that can deal with noise degradation, deformation of vector distortion, translation, rotation and scale invariant. The proposed system for the symbol recognition consists of 3 steps: 1) degradation model preprocessing step, 2) a different normalization for the second moment invariant and a measure for roundness and eccentricity for feature extraction step, 3) k-Nearest Neighbor with Mahalanobis distance compared to Euclidean distance and k-D tree for classifier. A comparison using multi-layer feed forward neural network classifier is given. An improvement of the discriminant factor around 4 times is achieved compared to that of the original normalized second moments using GREC 2005 dataset. Experimentally we tested our system with 3300 training images using k-NN classifier and on all 9450 images given in the dataset and achieved recognition rates higher than 86 % for all degradation models and 96 % for degradation models 1 to 4.

1 Introduction

The computational power of computer is increasing tremendously. This allows us to accomplish difficult tasks for pattern matching, symbol recognition, character recognition, finger print recognition, speech processing, etc. The moment invariants proposed by [3] is a well-known method for pattern recognition [4][5][6], but experimental results show a significant deviation value due to noise degradation, deformation, rotation and scale changes. Previous work [8] gave a revised fundamental theorem of moment invariants. Our work is inspired by [7] that gave the method of normalization to determine invariants. We proposed other normalization that yields a higher discriminant factor. The normalizer is derived from the property of SVD decomposition, which is the best linear unbiased estimator [4] for least square optimization. We define the roundness and eccentricity as a result of normalization, give

W. Liu and J. Lladós (Eds.): GREC 2005, LNCS 3926, pp. 408 – 421, 2006.

its correlation with the normalized second moment invariants [3] and evaluate our normalization using various classifiers such as k-nearest neighbor, k-D tree and neural network classifier to show its separability. In this paper we apply our proposed robust moment invariant for the symbol recognition and compare the results with [2]. Our solution is close to the previous work [1], which uses the statistical method with histogram of pixel-based descriptor and Manhattan distance. Our method to solve this problem does not use a histogram, but it is a statistical method based on binary image pixel position using three steps processing: image preprocessing, feature extraction and classification as discussed in Section 2. In Section 3 we discuss the experimental results and the conclusion is found in Section 4.

2 The Proposed Framework

2.1 Preprocessing Step

Noise Filtering and Degradation Level Measurement

The first step in our recognition system is a noise filtering and degradation level measurement. For an 8-bit image, a simple technique to measure the degradation level is by filtering the image with a low-pass filter kernel [1 4 6 4 1]/16 for both vertical and horizontal directions, and binarizing it with a threshold value t_h = 36/256. The threshold value range is around t_h (for pepper noise) to $2.5t_h$ (for hard pencil noise). The degradation level d_l is obtained from the number of pixels in the smooth area after filtering and dividing by the number of pixels in the original image.

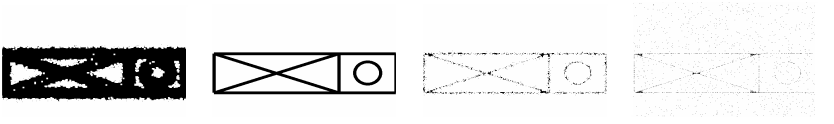
$$d_l = \frac{\sum_x \sum_y \left(f(x,y) * g(x,y,\sigma) > t_h \right)}{\sum_x \sum_y f(x,y)} \tag{1}$$

where $f(x,y)$ is the image and $g(x,y,\sigma)$ is the Gaussian low pass filter kernel with a 5 × 5 mask. This measurement allows us to detect how bad the noise level is. We used this measurement on GREC 2003 dataset [11]. For the worst degradation levels 7, 8, and 9, we have different preprocessing tasks. The degradation level value is around one for ideal-test and it will be higher for a noisy image. The highest degradation level appears when pixels in the noisy and thin lines merge after filtering.

This approach works quite well on GREC 2003 dataset. Using only one measurement, the degradation level measurement, we are able to separate all noise types and levels quite well. Unfortunately on GREC 2005 dataset [13], we are unable to separate noise type well enough using only degradation level measurement. So we combined degradation level measurement with line width measurement, noise distribution measurement and noise level measurement proposed in [12] to describe the noise features of each individual engineering drawing. With these measurements we have more degree of freedom to describe noise type and level in engineering drawings.

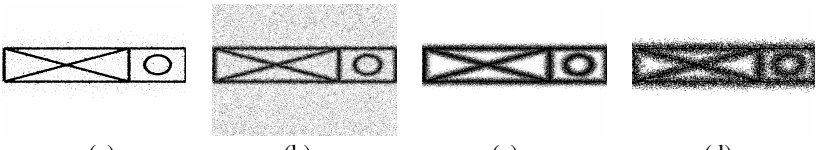
Based on the measurement results of primitives and noise we classify noise type and noise level of the drawing. From the combination of noise level measurement and ratio of primitives line width and image resolution we categorized the drawing into 4 categories, which are image with thick primitives (Fig. 1a), normal primitives (Fig. 1b), thin primitives (Fig. 1c), and very thin primitives (Fig. 1d).

Using a combination of noise distribution and degradation level measurement, we categorized noise type into 4 categories of noise type and distribution, which respectively are almost noiseless image (Fig. 2a), Gaussian noise combined with some high frequency noise (Fig. 2b), hard pencil noise concentrated around primitive (Fig. 2c), and hard pencil noise and high frequency noise combined with some Gaussian noise concentrated around the primitive (Fig. 2d).



(a) (b) (c) (d)

Fig. 1. Example of 4 categories of primitive thickness



(a) (b) (c) (d)

Fig. 2. Example of 4 categories of noise type and distribution

Morphological Operation

The next step of preprocessing is applying some morphological operations such as erosion, opening, closing, dilation, and thinning different categories of primitives and noise type. In this symbol recognition case, we experimentally categorize the images into 12 categories according to the primitives and noise type. Their corresponding image preprocessing tasks are shown in Table 1.

Table 1. The preprocessing task for different degradation levels and models

image type	preprocessing steps
I	open(0.25 lineW)→ close(0.4 lineW)→ idealizing line width→ thinning + clean
II	idealizing line width→ thinning + clean
IIA	close(0.25 lineW)→ open(0.3 lineW)→ idealizing line width→ thinning + clean
IIB	median filter(1.5 lineW)→ open(0.4 lineW)→ idealizing line width→ thinning + clean
IIC	erode(0.35 lineW)→ close(0.6 lineW)→ idealizing line width→ thinning + clean
IID	median filter(1.5 lineW)→ open(lineW)→ idealizing line width→ thinning + clean
III	close(0.25 lineW)→ open(0.3 lineW)→ idealizing line width→ thinning + clean
IIIA	close(2 lineW)→ dilate(0.1 lineW)→ close(lineW)→ idealizing linewidth→ thinning+clean
IIIB	median filter(1.5 lineW)→ open(0.25 lineW)→ idealizing line width→ thinning + clean
IIIC	median filter(1.5 lineW)→ close(0.25 lineW)→ open(0.3 lineW)→ idealizing line width→ thinning + clean
IVA	dilate(lineW)→ close(4 lineW)→ idealizing line width→ thinning + clean
IVB	filter→ dilate(1.5 lineW)→ close(4 lineW)→ component labelling→ remove small component→ idealizing line width→ thinning + clean

In the image type column of the above table, image types (I-IV) correspond to the thickness of image primitives, where (I) is thick primitive, (II) is normal primitive, (III) is thin primitive, and (IV) is very thin primitive, and (A-D) correspond to noise type and distribution. *LineW* is the line width of primitive and the diameter of disc structuring element used in open, close, erode, or dilate operation is shown in the bracket after each operation.

2.2 Feature Extraction

1. Roundness measurement

First, we propose a roundness measurement as the minimum singular value divided by the maximum singular value.

$$\text{roundness} = \sigma^2_{\min} / \sigma^2_{\max} \tag{2}$$

The roundness value is between 0 and 1. For example, a line has roundness equal to 0 with the smallest singular value equals to 0 and a disk has roundness equal to 1 with its singular values equal to identity. We compare this with the measurement for roundness or compactness defined by A.K. Jain [4] as follows:

$$\text{roundness} = \text{perimeter}^2/(4\pi \text{ area}) \tag{3}$$

Our measurement is more general as the object can be any shape. Whereas the equation (3) above may give value greater than 1 for a square shape and it depends on two parameters: the perimeter and area. The problem is that the measurement of area is not robust due to object deformation. Our proposed measurement is statistically robust. It gives the best estimation of roundness between 0 and 1, a higher discriminant factor and invariant to translation, rotation and scale changes. Next, we will show its relationship with the current moment invariant in Hu [3] and introduce a correction for normalization to achieve the same result as our proposed definition. Before that we will give a short explanation of singular values from the covariance matrix. A 2-dimensional unnormalized covariance matrix **X** is defined as follows:

$$\text{cov}(\mathbf{X}) = \mathbf{X}\mathbf{X}^{\mathrm{T}} = \begin{bmatrix} x-\bar{x} \\ y-\bar{y} \end{bmatrix} \begin{bmatrix} x-\bar{x} & y-\bar{y} \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \tag{4}$$

The singular values obtained from Singular Value Decomposition SVD are as follows:

$$\mathbf{U}\mathbf{D}\mathbf{V}^{\mathrm{T}} = \text{SVD}(\mathbf{X}\mathbf{X}^{\mathrm{T}}) \quad \text{and} \quad \mathbf{D} = \text{diag}(\sigma^2_{\max}, \sigma^2_{\min}) \tag{5}$$

$$\sigma^2_{\max} = \left| (a+c) + \sqrt{(a-c)^2 + 4b^2} \right| / 2 \tag{6}$$

$$\sigma^2_{\min} = \left| (a+c) - \sqrt{(a-c)^2 + 4b^2} \right| / 2 \tag{7}$$

where **D**, the diagonal matrix of singular values, is always positive and represents the variances of data distribution among its axes. The first singular value is the maximum singular value, which is the norm of covariance matrix. Next, we propose to use this norm as a normalizer of the second moments invariant [3]. The value of the square root will represent the eccentricity of data distribution. **U** and **V** are the orthogonal matrices that give the best rotation orientation along its principal axis. Now we review

the well-known feature in statistical-based pattern recognition, i.e., moment invariants. The central moments μ_{pq} and the normalized central moments μ_{pq}' in Hu [3] are defined as:

$$\mu_{pq} = \Sigma_x \Sigma_y \ (x - \overline{x} \ )^p \ (y - \overline{y} \ )^q f(x,y) \tag{8}$$

$$\mu_{pq}' = \mu_{pq} / \mu_{00}^{\gamma} \quad \text{and} \quad \gamma = (p+q+2) / 2 \tag{9}$$

We focus on two of the seven invariant moments, which are the second moments:

$$\phi_1 = \mu_{20} + \mu_{02} \tag{10}$$

$$\phi_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \tag{11}$$

The first equation ϕ_1 corresponds to roundness that gives 100 % correlation if and only if the normalized equation for the second moments is as follows:

$$\mu_{pq}' = \mu_{pq} / \sigma^2_{\max} \tag{12}$$

Next, the second equation ϕ_2 corresponds to eccentricity. It happens because $\mu_{20} = a$, $\mu_{02} = c$ and $\mu_{11} = b$. So the normalized first equation ϕ_1' equals to:

$$\phi_1' = (\mu_{20} + \mu_{02}) / \sigma^2_{\max} = (\sigma^2_{\max} + \sigma^2_{\min}) / \sigma^2_{\max} = 1 + \text{roundness} \tag{13}$$

and the normalized second equation ϕ_2' equals to:

$$\begin{aligned} \phi_2' &= ((\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2) / (\sigma^2_{\max})^2 = ((\sigma^2_{\max} - \sigma^2_{\min}) / \sigma^2_{\max})^2 \\ &= (1 - \text{roundness})^2 = \text{eccentricity}^2 \end{aligned} \tag{14}$$

Table 2 shows our experimental result with the symbol recognition database from GREC 2003 contest [11], which yielded clean images after using a simple preprocessing and GREC 2005 contest [13], which yielded noisier images even after using a complex preprocessing. The proposed normalized equations ϕ_1' and ϕ_2' have 5.2 and 2.6 times higher discriminant factor or 2.2 and 1.6 times higher recognition rate than the first original equations ϕ_1 and ϕ_2, respectively.

Table 2. The improvement of the proposed normalized second moments

No.	Parameter	ϕ_1'	ϕ_1	ϕ_1'/ϕ_1	ϕ_2'	ϕ_2	ϕ_2'/ϕ_2
1	Discriminant factor GREC03	22.8	3.2	7.1	30.5	6.1	5.0
2	Discriminant factor GREC05	21.2	5.2	4.1	24.5	14.9	1.6
	Average discriminant factor	22.0	4.2	5.2	27.5	10.5	2.6
1	Recognition rate GREC03	50.9 %	25.4 %	2.0	50.0 %	36.5 %	1.4
2	Recognition rate GREC05	21.2 %	6.9 %	3.1	19.0 %	7.6 %	2.5
	Average recognition rate	36.1 %	16.2 %	2.2	34.5 %	22.1 %	1.6

2. Radius min-max ratio

The radius min-max ratio is defined as the ratio between minimum radius and maximum radius $r_{\min}/r_{\max}$, where r is the Euclidean distance of each pixel to the centroid:

$$r = \sqrt{(x - \overline{x})^2 + (y - \overline{y})^2} \tag{15}$$

3. The compactness is defined as the perimeter in Eq. 19 divided by the bounding box area in the principal axis.

$$\text{compactness} = (\Sigma\, r) \ / \ ((y'_{\max} - y'_{\min}) \ (x'_{\max} - x'_{\min})) \tag{16}$$

where x' and y' are the rotated image in the principal axis and obtained from:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \mathbf{U}'^{\mathrm{T}} \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} \tag{17}$$

$$\mathbf{U}' = \mathrm{diag}(1, \det(\mathbf{U}))\, \mathbf{U} \tag{18}$$

where $\mathbf{U}$ is the rotation matrix correspond to the normalization of orientation to the principal axis and a correction $\mathbf{U}'$ is needed to remove the ambiquity since the singular value is always positive and the determinant of $\mathbf{U}$ may not always equal to one.
The perimeter is defined as the sum of all the distance of each pixel to the centroid:

$$\text{perimeter} = \int \sqrt{x^2(t) + y^2(t)}\; dt \equiv \Sigma\, r \tag{19}$$

where t is necessarily the boundary parameter but not necessarily its length.

4. Normalized pixel-perimeter is defined as the number of pixels, N, multiplied by the perimeter and divided by the maximum singular value.

$$\text{Normalized pixel-perimeter} = \mathrm{N}\,(\Sigma\, r) \,/\, \sigma^2_{\max} \tag{20}$$

5. Bounding box ratio $= (y'_{\max} - y'_{\min}) \,/\, (x'_{\max} - x'_{\min})$ (21)

6. Normalized perimeter is defined as the perimeter square divided by the number of pixels and the maximum singular value.

$$\text{Normalized perimeter} = \text{perimeter}^2 / (\mathrm{N}\sigma^2_{\max}) = (\Sigma\, r)^2 / (\mathrm{N}\sigma^2_{\max}) \tag{22}$$

7. Average standardized radius is defined as the mean of radii divided by its maximum radius.

$$\text{Average standardized radius} = (\Sigma\, r) \,/\, (\mathrm{N}\, r_{\max}) \tag{23}$$

8. Normalized perimeter square is defined as the perimeter square divided by the maximum singular value.

$$\text{Normalized perimeter square} = \text{perimeter}^2 / \sigma^2_{\max} = (\Sigma\, r)^2 / \sigma^2_{\max} \tag{24}$$

9. Inverse normalised perimeter is defined as the inverse of normalized perimeter (see Eq. 22).

$$\text{Inverse normalised perimeter} = (\mathrm{N}\sigma^2_{\max}) \,/\, (\Sigma\, r)^2 \tag{25}$$

10. Normalized second moment invariant for eccentricity is defined as:

$$\text{Eccentricity}^2 = ((\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2) \,/\, (\sigma^2_{\max})^2 \tag{26}$$

11. Average normalized angular is defined as the average of angular pixel distribution of every point around its centroid divided by its maximum value.

$$\text{Average normalized angular} = \Sigma\, \mathrm{P}(\theta) \,/\, (\mathrm{N_a} \max(\mathrm{P}(\theta))) \tag{27}$$

where $\mathrm{N_a}$ is the number of angular bin and $\mathrm{P}(\theta)$ is a probability density function of angular pixel distribution.

12. Average normalized radius is defined as the average of radial pixel distribution of every point around its centroid divided by its maximum value.

$$\text{Average normalized radius} = \Sigma\, \mathrm{P}(r) \,/\, (\mathrm{N_r} \max(\mathrm{P}(r))) \tag{28}$$

where $\mathrm{N_r}$ is the number of radial bin and $\mathrm{P}(r)$ is a probability density function of radial pixel distribution.

13. Average principal-axis-norm radius is defined as the mean of radii divided by the principal-axis length.

$$\text{Average principal-axis-norm radius} = (\Sigma\, r) \,/\, (\mathrm{N}\,(x'_{\max} - x'_{\min})) \tag{29}$$

2.3 Feature Selection

The feature selection is based on the discriminant factor d_f and its correlation to other features using Fisher discriminant analysis. The discriminant factor is obtained from the ratio between standard deviation "between class" σ_b and standard deviation "within class" σ_w. The first feature must have a high discriminant factor, and the next feature should have the lowest correlation toward zero as shown in Table 3.

Equation (30) until (33) is the corresponding equation for the discriminant factor and its variables. n_c is the number of classes which is 150 symbols from GREC 2005 contest which is 2.5 times the number of symbols in GREC 2003 contest. μ_b is the mean between classes which is the mean of the mean within class μ_{wj}.

Table 3. The comparison of discriminant factor, its correlations to the roundness and the recognition rate

No. feature	Feature name	GREC 2005 dataset					GREC 2003 dataset				
		σ_b	σ_w	d_f	cor. to no. 1	recog. rate %	σ_b	σ_w	d_f	cor. to no. 1	recog. rate %
1	roundness	.303	.0143	21.2	1.000	21.2	.299	.0131	22.8	1.000	50.9
2	radius min/max	.172	.0108	15.9	0.223	12.0	.180	.0122	14.8	0.193	34.0
3	compactness	.153	.0238	6.4	-0.147	8.8	.180	.0248	7.3	-0.287	36.0
4	pixelperimnorm	.189	.0228	8.3	0.767	9.1	.181	.0234	7.7	0.778	31.9
5	boundbox ratio	.237	.0123	19.3	0.948	16.6	.255	.0140	18.2	0.946	39.3
6	perimnorm	.159	.0077	20.5	0.976	19.6	.161	.0069	23.4	0.982	53.5
7	avgstdradius	.107	.0104	10.3	0.563	10.4	.089	.0095	9.4	0.621	26.1
8	perimsqrnorm	.133	.0348	3.8	0.823	7.3	.136	.0202	6.7	0.836	37.9
9	invperimnorm	.187	.0087	21.6	-0.946	19.9	.200	.0072	27.8	-0.941	53.6
10	eccentricity	.271	.0111	24.5	-0.966	19.1	.287	.0094	30.5	-0.968	50.0
11	avgangularbin	.172	.0241	7.1	0.471	7.4	-	-	-	-	-
12	avgradiibin	.081	.0254	3.2	-0.096	4.4	-	-	-	-	-
13	avgpanradius	.067	.0070	9.5	0.628	11.7	-	-	-	-	-

$$d_f = \sigma_b / \sigma_w \tag{30}$$

$$\sigma_b = (\Sigma_j\, (\mu_{wj} - \mu_b) \,/\, (n_c - 1))^{1/2} \tag{31}$$

$$\mu_b = (\Sigma_j\, \mu_{wj}) \,/\, n_c \tag{32}$$

$$\sigma_w = (\Sigma_j\, \sigma_{wj}) \,/\, n_c \tag{33}$$

2.4 Classifier

As a comparison, we use the simplest k-nearest neighbor (k-NN) classifier with Mahalanobis distance and compare its performance with Euclidean distance. To show the separability power of our robust moment invariant, we use k-D tree [9][10] classifier and multi-layer feed-forward neural network classifier [5] as discussed in Section 3. The average standard deviation within class σ_w is utilized as a weight to give a higher separability for a lower standard deviation among the features. A sample s is assigned

as class C_j if and only if the Mahalanobis distance d_M between the sample feature vector $\mathbf{s}_i$ and the mean within class μ_{wij} for all features i among classes j is minimum, where $i = 1\ldots n_f$ and $j = 1\ldots n_c$, where n_f is the number of features.

$$s \subset C_j \quad \text{iff} \quad \min_j d_M(\mu_{wj}, \mathbf{s}) = \min_j (\Sigma_i ((\mu_{wij} - \mathbf{s}_i)^2) / \sigma_{wi})^{1/2} \tag{34}$$

2.5 Training

Preprocessing results may not totally recover noisy images. To increase robustness in noisy images, we choose the image preprocessing with broken lines instead of continuous lines detection. Moreover, we also provide a filtering scheme in the training process of k-NN that removes the training images with unexpected groups of pixels. This selection process is based on a training error, which is the mean of distance between the feature vector of a sample image and a corresponding ideal or model image. In this case, we experimentally choose the threshold value of training error t_e to be around 0.0025 with the assumption that all features of element vector have values between 0 and 1. This will cover most of the good preprocessing results in the training set T to avoid mis-training with degrading images or over training such as the k-D tree result as shown in the Section 3.

$$s \subset T \quad \text{iff} \quad (\Sigma_i ((\mu_{wi} - \mathbf{s}_i)^2) / \sigma_{wi}) / n_f \; < \; t_e \tag{35}$$

3 Experimental Results

Using the simplest k-NN-based classifier, the best recognition rate can be achieved is around 99 % with a 7-feature vector (1-2-7-4-3-5-9) for GREC 2003 dataset and around 86 % with 8-feature vector (1-2-7-11-5-12-13-4) for GREC 2005 dataset as shown in Table 4. The highest recognition rate in GREC 2005 image dataset is achieved with a 7-feature vector by over training with degrading images. Hence, erroneous recognition may occur for the recognition of better preprocessing result, for

Table 4. The percentage of recognition rate vs the number of features

No. of features	Feature code	GREC 2005 image dataset			GREC 2003 image dataset		
		Ek-NN	Mk-NN	kD tree	Ek-NN	Mk-NN	kD tree
1	1	21.17	21.17	87.54	50.93	50.93	64.56
2	1-2	61.03	60.94	92.91	88.80	88.87	93.71
3	1-2-7	77.48	77.96	96.28	94.76	94.94	98.58
4a	1-2-7-4	78.41	79.53	95.51	97.78	98.25	99.71
5a	1-2-7-4-3	75.35	77.14	94.98	98.22	98.88	99.93
6a	1-2-7-4-3-5	78.17	79.82	95.79	98.47	98.98	99.93
7a	1-2-7-4-3-5-9	78.65	80.40	95.78	98.48	99.07	99.94
4b	1-2-7-11	80.84	81.97	96.20	-	-	-
5b	1-2-7-11-5	83.89	84.74	96.33	-	-	-
6b	1-2-7-11-5-12	84.13	85.50	96.32	-	-	-
7b	1-2-7-11-5-12-13	84.90	85.99	96.44	-	-	-
8	1-2-7-11-5-12-13-4	84.46	85.93	96.37	-	-	-

example from the ideal-test image dataset. This means that the subtotal testing recognition rate from the degradation model 1 to 4 may decrease as shown in Table 5. Similar problem may occur with other method such as k-D tree, but this is a kind of over training that does not represent the statistical-based classifier using Fisher discriminant analysis.

The recognition rate using Mahalanobis distance (Mk-NN) is 0.5 % to 3 % higher compared to those using Euclidean distance (Ek-NN) for the number of features greater than three. The highest recognition rate of 99.9 % for GREC 2003 dataset is obtained using k-D tree closest point with 2534 training images to construct a binary-tree classifier. The speed to construct a binary-tree classifier is very fast so that we can use more training set images. This highest recognition rate can be achieved since we incorporate the most extreme degradation, distortion, rotation and scale changes in the training set and it shows that our robust moment invariants have a higher separability power.

Table 5. The percentage of recognition rate distribution for GREC 2005 image dataset

		Robust Moment Invariant enhanced							RMI contest result as reference						
dataset	group	mod1	mod2	mod3	mod4	mod5	mod6	total	mod1	mod2	mod3	mod4	mod5	mod6	total
grec05	100	100	100	93	100	85	62	90.0	98	93	98	96	67	56	84.7
	150	100	100	98	98	87	56	89.8	98	87	98	93	58	56	81.7
	25	100	100	88	100	92	58	89.7	100	100	100	100	88	70	93.0
	50	100	100	100	100	96	60	92.7	98	94	100	100	72	66	88.3
grec05-rot+scl	100	82	94	98	74	64	42	75.7	78	76	50	38	48	8	49.7
	150	86	96	90	72	66	40	75.0	80	66	64	26	30	12	46.3
	25	82	92	98	60	64	42	73.0	78	82	40	28	44	12	47.3
	50	84	96	96	74	68	48	77.7	74	80	56	40	44	8	50.3
grec05-Rot	100	99	99	95	99	87	36	85.8	97	92	98	92	64	41	80.7
	150	99	100	96	97	84	29	84.2	94	91	96	85	59	24	74.8
	25	96	100	100	100	92	40	88.0	100	100	98	98	76	36	84.7
	50	98	98	100	100	90	44	88.3	92	96	98	96	70	40	82.0
grec05-Scl	100	92	94	96	84	68	56	81.7	64	66	78	42	54	18	53.7
	150	96	96	90	74	76	54	81.0	82	66	70	44	48	20	55.0
	25	92	100	94	78	80	60	84.0	80	82	60	38	56	20	56.0
	50	98	98	98	82	88	58	87.0	72	74	66	50	58	14	55.7
subtotal testing	A	95.1	98.1	95.6	89.3	81.5	48.4	84.7	88.6	85.4	83.0	71.6	59.2	33.9	70.3
sample-test-rot	100	99	100	95	98	89	27	84.7	-	-	-	-	-	-	-
	150	99	100	97	95	85	25	83.5	-	-	-	-	-	-	-
	25	100	100	100	100	72	28	83.3	-	-	-	-	-	-	-
	50	98	98	98	100	90	18	83.7	-	-	-	-	-	-	-
sample-test	100	100	100	96	100	81	66	90.5	-	-	-	-	-	-	-
	150	100	99	97	100	87	66	91.5	-	-	-	-	-	-	-
	25	100	100	96	92	88	68	90.7	-	-	-	-	-	-	-
	50	100	100	100	96	96	62	92.3	-	-	-	-	-	-	-
subtotal training	B	99.5	99.6	96.9	98.0	86.4	45.1	87.6	-	-	-	-	-	-	-
total dataset	A+B	96.6	98.6	96.1	92.4	83.2	47.2	85.7	-	-	-	-	-	-	-

Table 6. The percentage of performance class comparison with GREC 2005 contest results

dataset	RMI enhanced			RMI contest result			ZW contest result			FM contest result		
	mod 1-4	mod 5-6	total	mod 1-4	mod 5-6	total	mod 1-4	mod 5-6	total	mod 1-4	mod 5-6	total
grec05	98.6	73.8	90.3	96.4	64.2	85.7	97.8	75.2	90.3	91.9	92.3	92.1
grec05-rot+scl	85.9	54.3	75.3	59.8	25.8	48.4	88.5	48.5	75.2	82.1	58.3	74.2
grec05-rot	98.3	61.5	86.1	94.5	49.8	79.6	95.8	51.2	80.9	87.6	74.5	83.2
grec05-scl	91.4	67.5	83.4	64.6	36.0	55.1	90.3	66.0	82.2	81.0	76.8	79.6
subtotal testing	94.5	65.0	84.7	82.2	46.6	70.3	93.8	60.8	82.8	86.5	77.1	83.3
sampletest-rot	98.9	55.5	83.9	-	-	-	-	-	-	-	-	-
sampletest	98.1	76.0	91.2	-	-	-	-	-	-	-	-	-
subtotal training	98.5	65.7	87.6	-	-	-	-	-	-	-	-	-
total	95.9	65.2	85.7	-	-	-	-	-	-	-	-	-

The GREC 2005 image dataset consists of 150 symbols as shown in Figure 3 and total 9450 images which consist of 150 ideal-test images, 3300 training images (sample test set), and 6000 testing images. Whereas, the GREC 2003 image dataset is a sub-set of GREC 2005 image dataset. The GREC 2005 dataset is designed to test the scalability of symbol recognition system to cope with the increasing number of symbols with more severe degradation levels, but less degradation number of models. The corresponding results of the preprocessing step for the various degradation models are shown in Figure 4. We found that the feature vector for GREC 2003 dataset, which mainly based on the radii measurement, is not sensitive to the angular pixel variation

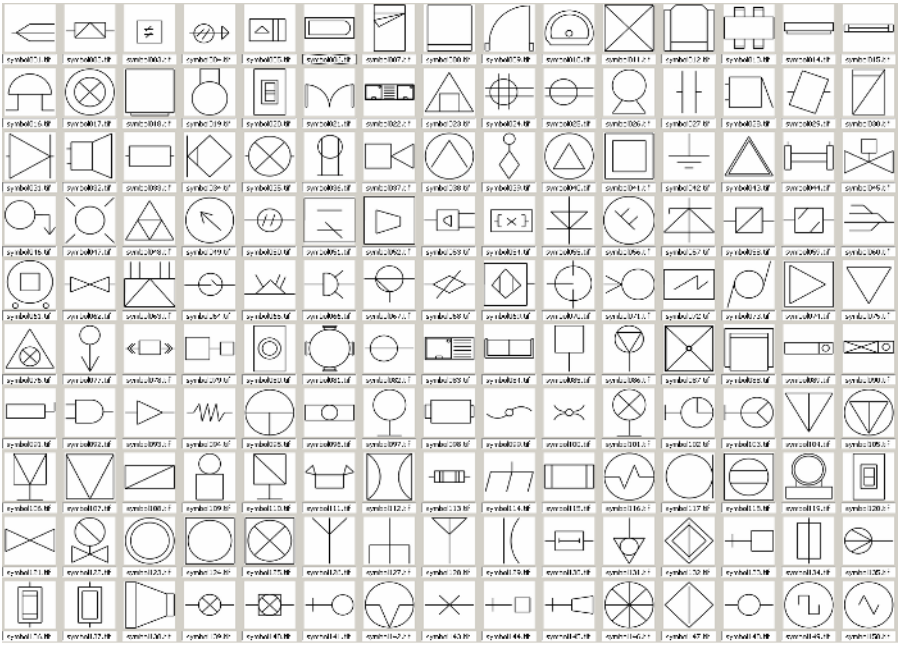


Fig. 3. The 150 symbols in GREC 2005 image dataset

that exists in GREC 2005 dataset. This case was not detected in the GREC 2005 contest since we did not have better noise preprocessing which resulted in a 70 % recognition rate. Adopting the Adaptive Noise Reduction preprocessing [12] yields 10 % improvement. To increase the separability power for better recognition rate, therefore, we have to add some additional features such as average normalized angular based on angular pixel distribution (feature number code 11) and radial pixel distribution to increase radii sensitivity (feature number 12 and 13). These three features yield 6 % additional gain up to 86 %.

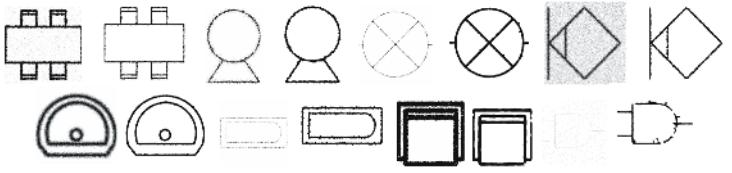


Fig. 4. The preprocessing results for the various degradation models degrad 4-7-9 & mod 2 to 6

The system has false alarm detection on the similar objects as shown in Figure 5 for symbol 11-87, 98-115, 20-120, 136-137, 149-150-49, 43-75, 48-23.

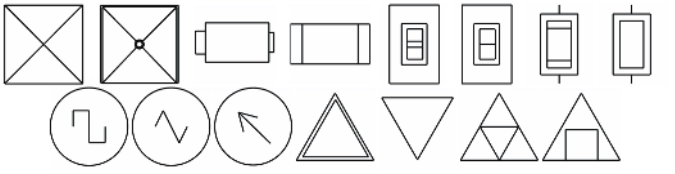


Fig. 5. False alarm detection on the similar objects for GREC 2005 dataset

Figure 6 shows the result of GREC 2003 dataset with the most difficult test images to be recognized are those with degradation 7 and scale-rotation. It also shows that our robust moment invariants are 100 % scale invariant, but it is slightly dependent on rotation changes since the orientation measurement used as the normalization in the principal axis is not as stable as the scale from the diagonal matrix of SVD.

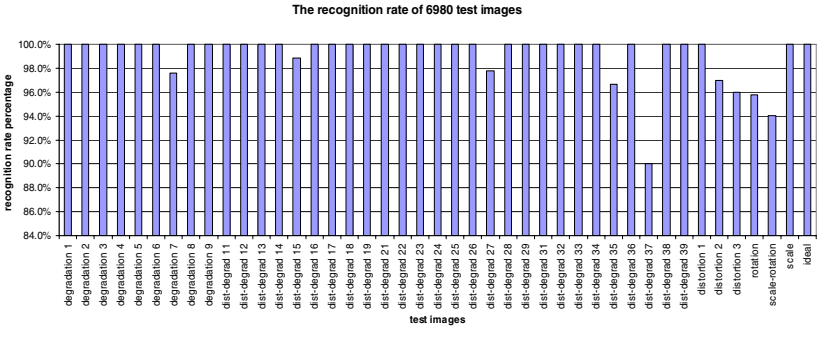


Fig. 6. The recognition result of 6980 images using Mk-NN classifier for GREC 2003 dataset

Figure 7 shows the result of GREC 2003 dataset that the symbols 1, 2, 4, 11, 47, 48 and 51 have adjacent distance such as 11-48-33, 51-15, 47-42, 2-5, 4-3 and 1-35, or severe degradation and distortion level. These happen since we use the centroid of each class to calculate the distance from the sample. However, using k-D tree classifier these cases do not happen since the constructed tree is trained using all the supporting vectors. It shows that all symbols can be separated well using k-D tree and our proposed moment invariant features.

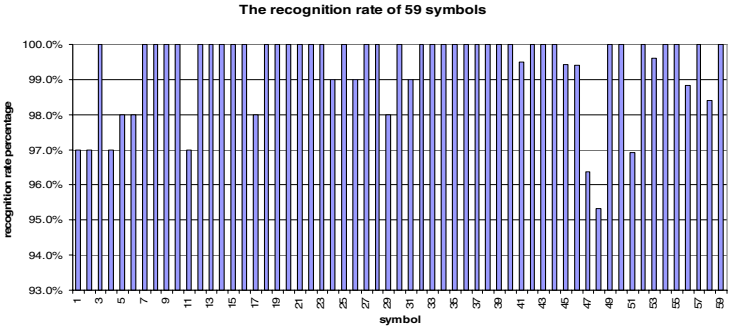


Fig. 7. The recognition result of 59 symbols using Mk-NN classifier for GREC 2003 dataset

A comparison result has also been done for multi-layer feed-forward neural network with a lower recognition rate of 91.48 % using 3 layers, 15:30:59 neurons, 59 ideal-test training set images, tangent sigmoid activation function for the hidden layer, pure linear activation function for the output layer and Levenberg-Marquadt optimization. This lower recognition rate happens because zero degradation, distortion, rotation and scale test images are trained with a tradeoff between the convergent speed and memory limitation. Figure 8 shows the result of GREC 2003 dataset that the convergence rate of the network training is relatively fast using a small training set and

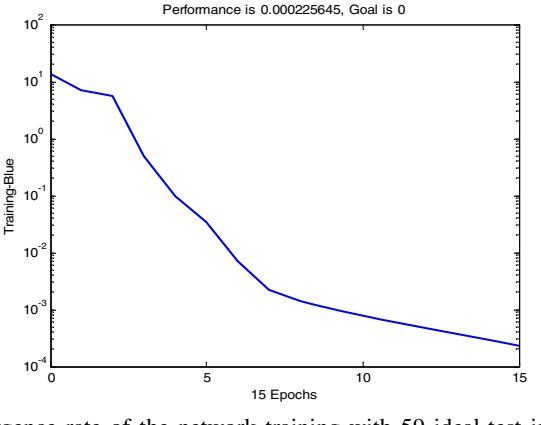


Fig. 8. The convergence rate of the network training with 59 ideal-test images and 10:30:59 neurons using Levenberg-Marquadt optimization for GREC 2003 dataset

achieved a Mean Square Error less than 0.001. It means that the separability power of all features is strong. The separability power based on the roundness is also shown in [14] for the first four features value: roundness, radius min/max, average standardized radius and normalized pixel-perimeter. The symbols that have the same values on 2 features would not have the same values for another feature. Most of the symbol can be separated using these 4 features with the recognition rate of 99 % as shown in Table 4.

4 Conclusion

In this paper, we propose novel robust moment invariants by normalizing the second moments using the norm of covariance matrix and defining the roundness and eccentricity measurement to yield a higher discriminant factor of around 5 times by average the test results on GREC 2003 and GREC 2005 datasets. We experimentally evaluate the proposed framework using various classifiers such as k-NN based on Fisher discriminant analysis, k-D tree and neural network and show that by using Mahalanobis k-NN, higher separability with improved recognition rate can be achieved, which is around 86 % for GREC 2005 dataset and 99 % for GREC 2003 dataset. Note that our feature vector is still maintained to be less than 10 features, which is useful for any object recognition based on statistical and geometrical analysis.

5 Discussion

Our strategies to achieve general symbol recognition are: First, to deal with various noises processing model is to classify noise model automatically in some measurements with smooth step changing value in various model degradations. The more noise measurement model we have, the better noise preprocessing model we get. For real implementation, other problem will happen, i.e., the localization of symbols as the next challenge. Second, the roundness measurement is inspired by the scale space problem, the properties of singular value decomposition and invariant to scale and rotation. Basically the proposed roundness is a dimensionless measurement for the second moment or covariance measurement in x-y axes, which will then be transformed into principal/major axis using SVD. Singular value itself is a stable measure regardless of the orientation, but it still has a dimension. The maximum and minimum singular values are the covariance along major and minor axis respectively. The normalization of minimum singular value by maximum singular value is performed so that the dimension can be eliminated and the measurement always stable between 0-1.

References

1. Yang, S.: Symbol Recognition via Statistical Integration of Pixel-Level Constraint Histograms: A New Descriptor. IEEE Trans Pattern Analysis & Machine Intelligent, vol. 27, no. 2, (Feb. 2005) 278-281.
2. Valveny, E., and Dosch, P.: Symbol Recognition Contest: A Synthesis. In: Llados, J. and Kwon, Y.B. (eds.): Graphics Recognition: Recent Advances and Perspectives, GREC 2003. Lecture Notes in Computer Science, vol. 3088. Springer-Verlag, Berlin Heidelberg New York (2004) 368-385.

3. Hu, M.K.: Visual Pattern Recognition by Moment Invariants. IRE Trans. Information Theory, vol. 8, (1962) 179-187.
4. Jain, A.K.: Fundamentals of Digital Image Processing. Prentice-Hall Inc. (1989).
5. Gonzales, R.C. and Woods, R.E.: Digital Image Processing, Prentice-Hall Inc. (2002).
6. Mindru, F., Tuytelaars, T., Gool, L.V., and Moons, T.: Moment Invariants for Recognition under Changing Viewpoint and Illumination. Computer Vision and Image Understanding, vol. 94, (2004) 3-27.
7. Rothe, I., Susse, H., and Voss, K.: The Method of Normalization to Determine Invariants. IEEE Trans. Pattern Analysis and Machine Intelligent, vol. 18, no. 4, (Apr. 1996) 366-376.
8. Reiss, T.H.: The Revised Fundamental Theorem of Moment Invariants. IEEE Trans. Pattern Analysis and Machine Intelligent, vol. 13, no. 8, (Aug. 1991) 830-834.
9. DeBerg, M., vanKreveld, M., Overmars, M., and Schwarzkopf, O.: Computational Geometry: Algorithms and Applications, Springer Verlag, (2000).
10. http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=7030&objectType=file
11. http://www.cvc.uab.es/GREC2003/SymRecContest/index.htm
12. Zhang, J., Zhang, W., and Wenyin, L.: Adaptive Noise Reduction for Engineering Drawings based on Primitives and Noise Assessment. Sixth IAPR Int. Workshop on Graphics Recognition GREC 2005, (Aug. 2005) 136-145.
13. http://symbcontestgrec05.loria.fr/index.php
14. Weliamto, W.A., and Seah, H.S.: Robust Moment Invariant with Higher Discriminant Factor for Symbol Recognition. Sixth IAPR Int. Workshop on Graphics Recognition GREC 2005, (Aug. 2005) 68-76.