Comparison Between K-Nearest Neighbors, Self-organizing Maps and Optimum-Path Forest in the Recognition of Packages using Image Analysis by Zernike Moments

Rodrigo Dalvit Carvalho da Silva, David Nascimento Coelho, George André Pereira Thé, Marcel Ribeiro Mendonça
Reference Centre for Automation and Robotics
Department of Teleinformatics Engineering
Federal University of Ceará, Fortaleza, Brazil
rodrigodalvit@hotmail.com, davidcoelho89@gmail.com, george.the@ufc.br, marcelribeiroceara@gmail.com

***Abstract*—Recognition of objects using an industrial image sensor is an important tool that has been motivated by the necessity of automatic recognition systems in the industrial automation. In this context, an interesting problem is the automatic image acquiring and a high reliability in objects classification. To this end, this paper presents a comparison between k-Nearest Neighbors Classifier using Euclidean, City Block, Cosine and Correlation distance metric, the Self-Organizing Map (SOM) - Artificial Neural Network (ANN) and the Optimum-Path Forest, for classification of images taken from a low-resolution industrial sensor. Classification performance has been compared in terms of extraction time and accuracy using image analysis by Zernike moments.**

1. INTRODUCTION

Machine vision brings innovative solutions to industrial applications. A plethora of industrial activities have been benefited from the use of this technology on manufacturing processes, because it improves productivity and quality management providing a competitive advantage to industries [1].

Thanks to recent developments in data acquisition, processing, and process control systems, efficiency of many of the industrial applications has been improved with the automated visual processing and classification systems help. Technological advances in digital image acquisition and processing have allowed automated visual inspection systems build [2]. However, even with the rise of high-capacity computers, classification between objects has proved to be a complex problem for machines.

In this context, this paper describes a comparison performance in terms of efficiency and processing time between the classifiers k-NN, Self-Organizing Maps - Artificial Neural Network and the Optimum-Path Forest aimed at improving the ranking process goals using images taken from a sensor with low resolution.

The remainder of the paper is directed as follows. Section 2 describes the feature extraction process. Section 3 gives an overview of supervised classification method. Section 4 discusses the results. Finally, section 6 concludes the paper with a brief discussion of future research.

2. FEATURE EXTRACTION PROCESS

Moments and moment functions have been extensively used for feature extraction in pattern recognition and object classification. One important property of the moments is their invariance under affine transformation.

Moments are scalar quantities used to characterize a function and to capture its significant features. From the mathematical point of view, moments are projections of a function onto a polynomial basis [3].

M. K. Hu [4] introduced the concept of moment, since than invariant moments and moment functions have been widely used in the fields of image analysis and pattern recognition. Hu's moment descriptors are invariant with respect to scale, translation and rotation of the image. However, the kernel function of geometric moments of order $(p + q)$, is not orthogonal, thus the geometric moments suffer from the high degree of information redundancy, and they are sensitive to noise for higher-order moments [5].

Zernike moments were first used in image analysis by Teague [6], which proposed this moments based on the basis set of orthogonal Zernike polynomials. Zernike moments are the mapping of an image onto a set of complex Zernike polynomials. As these Zernike polynomials are orthogonal to each other, Zernike moments can represent the properties of an image with no redundancy or overlap of information between the moments [7].

978-1-4799-5551-0/14/$31.00 ©2014 IEEE

2.1 Zernike Moments

F. Zernike [8] introduced a set of complex polynomials which form a complete orthogonal set over the interior of the unit circle, $x^2 + y^2 = 1$, where $Vnm(\rho, \theta) = Vnm(x, y)$ and,

$$V_{nm}(x, y) = R_{nm}(\rho) exp(jm\theta) \quad (1)$$

where $j = \sqrt{-1}$, n is a positive integer or zero, m is a positive and negative integers subject to constraints $n - |m|$ even, $|m| \leq n$, ρ is the length of vector from origin to (x,y) pixel

$$\rho_{xy} = \frac{\sqrt{(2x - N + 1)^2 + (N - 1 - 2y)^2}}{N} \quad (2)$$

θ is the angle between vector ρ and x axis in counterclockwise direction

$$\theta_{xy} = \tan^{-1} \frac{N - 1 - 2y}{2x - N + 1} \quad (3)$$

$R_{nm}(\rho)$ is the radial polynomial defined as

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} c(n, m, s) \rho^{n-2s} \quad (4)$$

and

$$c(n, m, s) = (-1)^s \frac{(n - s)!}{s! \left(\frac{n + |m|}{2 - s}\right)! \left(\frac{n - |m|}{2 - s}\right)!} \quad (5)$$

The discrete form of the Zernike moments of an image size N×N is expressed as follows [9],

$$Z_{nm} = \frac{n + 1}{\lambda_N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} V_{nm}(x, y) I_{xy} \quad (6)$$

A plot of the polynomial values of order 0-5 with low repetitions, obtained from Eq. (4) is given in Fig. 1.

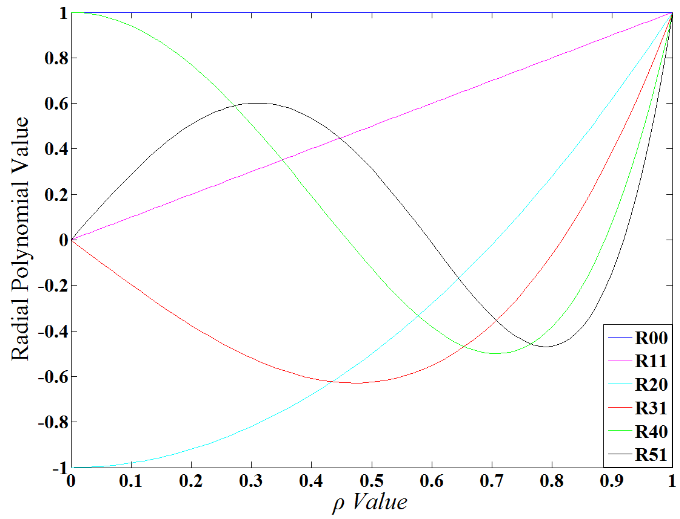

Figure 1. Six orthogonal Zernike radial polynomial plots.

3. SUPERVISED CLASSIFICATION

The final stage of any image-processing system where each unknown pattern is assigned to a category is the classification. The difficulty of the classification problem depends on the variability in the characteristic values of the objects of the same class.

Let's suppose that we have a classification problem in which there are M possible classes and there are N independent and identically distributed samples Z. The supervised classification problem consists in using that prior knowledge to classify new samples Zs to one of the M possible classes in a manner to minimize the classification error.

3.1 k-Nearest Neighbors Classifier

The k-Nearest Neighbor rule (k-NN) is one of the best known methods for supervised pattern recognition in analytical chemistry and, more generally, the method has been proposed by T. M. Cover [10] as a reference method for the evaluation of the performance of more sophisticated techniques [11].

In general the following steps are performed for k-NN algorithm:

- Choose of k value: k value is completely up to user. Generally after some trials a k value is chosen according to results.
- Distance calculation: Any distance measurement can be used for this step.
- Distance sort in ascending order: Chosen k value is also important in this step. Found distances are sorted in ascending order and k of minimum distances are taken.
- Classification of nearest neighbors: Classes of k nearest neighbor are identified.

As shown in step 2, we can use various distances.

3.1.1 Euclidean Distance

Let's considerer the two input variable case since it is easy to represent in two-dimensional space. The distance between these two points is computed as the length of the difference points, is denoted by

$$d(x,x') = |x - x'| = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2} \tag{7}$$

3.1.2 City Block Distance

The City Block distance between two points, x and x', with k dimensions is calculated as

$$d(x,x') = \sum_{j=1}^{k} |x_j - x'_j| \tag{8}$$

3.1.3 Cosine Distance

The Cosine distance between two points

$$d(x,x') = 1 - \frac{\sum_{i=0}^{k}(x_i x'_i)}{\sqrt{\sum_{i=0}^{k}(x_i)^2 \sum_{i=0}^{k}(x'_i)^2}} \tag{9}$$

3.1.4 Correlation Distance

The Correlation distance between two points

$$d(X,X') = 1 - \frac{\sum_{i=0}^{k}(X_i - \overline{X}_i)(X'_i - \overline{X}'_i)}{\sqrt{\sum_{i=0}^{k}(X_i - \overline{X}_i)^2 \sum_{i=0}^{k}(X'_i - \overline{X}'_i)^2}} \tag{10}$$

3.2 Optimum-Path Forest (OPF)

J. P. Papa et al. [12] introduced the idea of designing pattern classifiers based on optimum-path forest that was developed as a generalization of the Image Foresting Transform (IFT) [13]. OPF is simple, multi-class, parameter independent, does not make any prior assumption about the shapes of classes and can handle some degree of overlapping [14].

This classifier is based on a forest of optimal paths, which is constructed by calculating the maximum of the shortest paths between the samples and the prototypes of classes. The path cost is computed from the distances between the feature vectors of the samples.

The OPF is divided into two steps, adjustment and prediction. In step adjustment is made learning classifier. The main components of OPF are calculated: Minimum Spanning Tree, prototypes and the cost matrix samples compared to the prototypes.

In step Prediction, new samples are classified using the forest paths resulting from the great stage of adjustment.

The OPF classifier reduces the problem of classification standards for a partitioning problem in a graph. The prototypes chosen initially begin a process of conquest in the graph, offering optimal cost paths to the other samples. A path cost function is defined, which associates with each path in the graph the cost of considering all objects along the way as belonging to the same class. Thus, the graph is partitioned into a forest of optimal paths whose roots are the prototypes. Prototypes competing, each wave defines a tree of optimal paths.

The training consists essentially in building this great forest paths where the objects in a given OPF will have the same label its prototype, in other words, the same class from the root of the tree of which he belongs. To classify an object in the training set, we evaluate the optimal paths from prototypes to him in order to find which OPF will win the element to be classified. The label of this tree is associated with the test object.

3.3 Self-Organizing Map (SOM)

The Self-Organizing Map (SOM) [15], is an unsupervised neural network that has the ability to perform clustering and preserve the topology [16] of the data.

The general equation utilized, at this paper, to update the neurons is as follows [17]

$$w_i(t+1) = w_i(t) + \eta(t)h(i^*,i,t)[x(t) - w_i(t)] \tag{11}$$

in which $w_i(t+1)$ is the new neuron weight, $\eta(t)$ is the learning rate and $h(i^*,i,t)$ is the neighborhood function. The learning rate varies according to

$$\eta(t) = n_0\left(1 - \frac{t}{t_{max}}\right) \tag{12}$$

in which $0 < n_0 < 1$, t is the current iteration and t_{max} is the total number of iterations.

The neighborhood function varied according to:

$$h(i^*,i,t) = exp\left(-\frac{\|r_i(t) - r_{i*}(t)\|^2}{2\alpha^2(t)}\right) \tag{13}$$

in which i^* is the current winner neuron, $\|r_i(t) - r_{i*}(t)\|$ is the squared Euclidian distance between the current neuron and the current winner neuron, and α^2 determines the influence of the winner neuron over the others.

This neural network can also be used as a classifier.

At the end of its training, the data are shown to the network again, and each neuron is labeled as a representative of some class. The choice of the class to which the neuron will be labeled is by counting the number of data of this class, for which the neuron was the winner.

4. Results and Discussion

The hardware used for image acquisition in this work was the 50x64 resolution 3D sensor effector pmd E3D200, from ifm electronic®. It contains Ethernet interface, thus allowing for implementation of remote and eventually real-time applications of classification algorithms and, eventually, real-time. The device has been used to acquire images of three packages with a few differences in size, as shown in Fig. 2, 3 and 4.

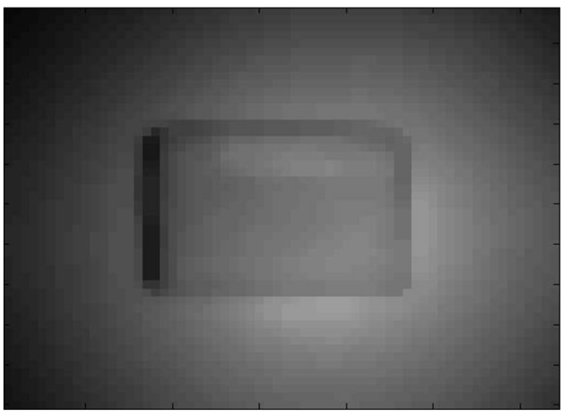

Figure 2. Package 1 with dimension 15×10.5×7.2 cm.

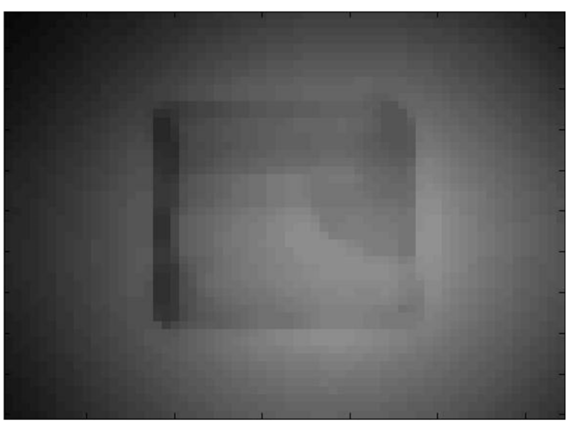

Figure 3. Package 2 with dimension 15×14×6 cm.

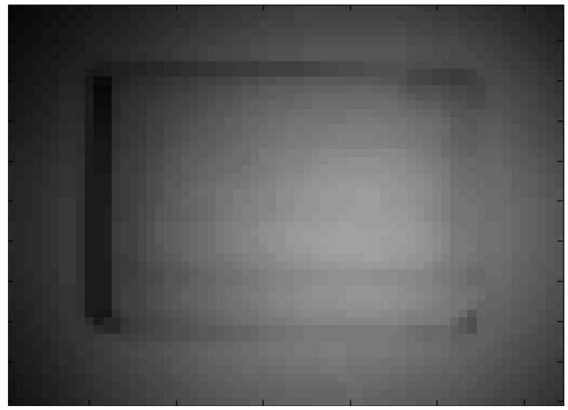

Figure 4. Package 3 with dimension 21.5×16.2×9.6 cm.

It is worth lighting that the experiments are based on three classes. The number of prototypes per class is 6, referring to the 6 sides of each box, so, the database contains 18 objects.

In our experiments, the number of input features extracted and the processing times per 50 times is shown in Table 1.

TABLE 1. Number of input × Processing Times of the Zernike moment.

Number of Input	Times (seconds)				
	Min	Max	Mean	Median	STD
36	0.360	0.7806	0.4607	0.452	0.446

These input vectors are presented to the classifiers. The classifiers was trained and tested 150 times with the same database. The experimental results showed the k-NN classifier, are given in Table 2.

TABLE 2. Accuracy of the k-NN classifier.

k-NN (distance)	Accuracy (%)				
	Min	Max	Mean	Median	STD
Eclidean	70.00	97.33	81.63	79.00	10.18
CityBlock	70.00	**98.00**	80.19	77.00	10.38
Correlation	70.00	96.00	**83.19**	**85.33**	10.18
Cosine	**71.33**	96.67	82.59	83.33	9.28

Among the k-NN classifier, the one based on the City Block distance obtained the best hit rates for variations in kbetween 1-18. However, the Correlation distance obtained the best rates mean and median.

As the data used were taken from an industrial sensor, which is used in industrial processes in real time, it becomes important to determine the time required for extraction, which was presented in Table 1, and classification. Table 3 shows the time required for classification of k-NN classifier.

TABLE 3. PROCESSING TIMES FOR CLASSIFICATION.

k-NN (distance)	Times (seconds)				
	Min	Max	Mean	Median	STD
Eclidean	0.0058	0.0186	0.0100	0.0093	0.0002
CityBlock	0.0056	0.0194	0.0100	0.0093	0.0003
Correlation	0.0022	0.0187	0.0066	0.0056	0.0005
Cosine	0.0021	0.0171	0.0063	0.0053	0.0004

Tables 4, 5 and 6 shows the training times, times for classification and accuracy values, respectively, of OPF classifier and the SOM-ANN classifier.

TABLE 4. PROCESSING TIMES TO THE TRAINING.

	Times (seconds)				
	Min	Max	Mean	Median	STD
OPF	0.000039	0.00013	0.000087	0.000091	0.00002
SOM ANN	0.1976	0.2773	0.2120	0.2092	0.06000

TABLE 5. TABLE 5 PROCESSING TIMES TO THE CLASSIFY

	Times (seconds)				
	Min	Max	Mean	Median	STD
OPF	0.00016	0.0006	0.00038	0.0004	0.00008
SOM ANN	0.0052	0.0086	0.0059	0.0057	0.0013

TABLE 6. TABLE 6 ACCURACY

	Accuracy (%)				
	Min	Max	Mean	Median	STD
OPF	**94.00**	**98.00**	**96.42**	**97.00**	1.006
SOM ANN	71.33	91.33	82.23	80.67	4.26

For the proposed problem in this work, all classifiers proved compatible with the classification time required in the process of industrial image classification in real time.

The classifier based on SOM network had more time for classification, this is due to the fact this has more parameters than the classifier based on graph, OPF, which is free of parameters.

This neural network also presented lower hit rates, mainly due to the fact the first 2 classes having similar values.

A possible difficulty encountered by the classifier SOM is due to the fact that there are few training data. Unable to generate new training data because this feature extraction, Zernike Moment, is invariant to rotation, scaling and translation. Thus, were generated only 6 training data for each class, referring to the sides of each box.

On the other hand, even with these few training data, the OPF classifier achieved a maximum rating of 98% and a classification and training time small for the application in question. Furthermore, this classifier obtained a good average hit rate, 96.42%, and a median good hit rate, 97.00%.

5. CONCLUSIONS

This paper introduced a comparative study between 3 classifiers methods k-NN using Euclidean, City Block, Cosine and Correlation distance, Self Organizing-Map - ANN and Optimum-Path Forest to the recognize the images of an industrial sensor using Zernike moment to feature extraction.

At this writing, the best results with this technique of feature extraction has been through the OPF classifier, which proved to be fast and efficient for the problem at hand.

A second alternative would be to use the k-NN classifier, based on City Block distance, which proved an accuracy rate of 98.00% and a minimum time for training and classification. Another fact that makes it attractive for application is the extensive literature found on it and its easy implementation.

ACKNOWLEDGMENT

The institution research support CAPES for financial support. FUNCAP has supported this work (PP1-0033-00032.01.00/10).

Authors thank Rockwell Automation do Brasil, for the support through the Scientific and Technical Cooperation Agreement with the Federal University of Ceará. Authors thank NUTEC, for administrative facilities.

REFERENCES

[1] E. N. Malamas, E. G. M. Petrakis, M. Zervakis, L. Petit, J. D. Legat. A survey on industrial vision systems, applications and tools, Image and Vision Computing, 2003, vol. 21, pp. 171-188.

[2] M. A. Selver, O. Akay, F. Alim, S. Bardakçi, M. Olmez. An automated industrial conveyor belt system using image processing and hierarchical clustering for classifying marble slabs. Robotics and Computer-Integrated manufacturing, 2011, vol. 27, pp. 164-176.

[3] J. Flusser, T. Suk, B. Zitova, Moments and Moment Invariants in Pattern Recognition, Wiley, Chichester, 2009.

[4] M. K. Hu. Visual Pattern Recognition by Moment Invariants. IRE, Transactions on information theory, 1962, pp. 179-187.

[5] D. Sridhar, Dr I. V. M. Krishna. Face Recognition using Tchebichef Moments. International Journal of Information &Network Security, 2012, vol. 1, no. 4, October, pp. 243-254.

[6] M. Teague. Image analysis via the general theory of moments, J. Opt Soc. Am. 1980, vol. 70, no. 8, pp. 920-930.

[7] Sun-Kyoo Hwang and Whoi-Yul Kim, A novel approach to the fast computation of Zernike moments, Pattern Recognition, 2006, vol. 39, pp. 2065-2076.

[8] F. Zernike, Physica, 1934, vol. 1, pp. 689.

[9] S. K. Hwang, W. Y. Kim. A novel approach to the fast computation of Zernike moments. Pattern Recognition, 2006, vol. 39, pp. 2065-2076.

[10] T. M. Cover. IEEE Trans. Inf. Theory, 1968, IT-14 50.

[11] D. Coomans, D. L. Massart. Alternative k-Nearest Neighbor Rules in Supervised Pattern Recognition. Part. 1. k-nearest neighbor

classification using alternative voting rules. Anal. Chim. Acta., 1981, 136, 15.

[12] J. P. Papa, A. X. Falcão, C. T. N. Suzuki, (2009). Supervised pattern classification based on optimum-path forest. Int J Imaging Systems and Technology, 2009, vol. 19, no. 2, pp. 120-131.

[13] A. X. Falcão, J. Stolfi, and R. A. Lotufo. The image foresting transform: Theory, algorithms, and applications, IEEE Trans. on Pattern Analysis and Machine Intelligence, 2004, vol. 26, no. 1, pp. 19–29.

[14] R. Souza, R. Lotufo, L. Rittner. A Comparison between Optimum-Path Forest and K-Nearest Neighbors Classifiers. 25th Conference Graphics, Patterns and Images (SIBGRAPI) 2012, Aug. 260-267.

[15] T. Kohonen. Self-Organizing Map. Proceedings of the IEEE, 1990, vol. 78, no. 9.

[16] S. Wu, T. W. S. Chow. Induction machine fault detection using SOM-based RBF neural networks. IEEE Transactions ont Industrial Electronics, 2004, vol. 51, no. 1.

[17] S. Haykin. Redes Neurais: Princípios e Práticas. Porto Alegre: Ed. Bookman, 2001