Moment Invariant based Classification of objects from low-resolution Industrial Sensor Images

Rodrigo D. C. Silva

Department of Teleinformatics Engineering
Federal University of Ceará
Fortaleza, Ceará, Brazil
rodrigodalvit@hotmail.com

George A. P. Thé

Department of Teleinformatics Engineering
Federal University of Ceará
Fortaleza, Ceará, Brazil
george.the@ufc.br

***Abstract* — In this paper, the issue of object recognition for industrial applications, using images extracted from a 3D sensor is discussed. We focus on moment invariants based feature extraction algorithms for the classification of images from an industrial 3D sensor, using the Optimum-Path Forest (OPF) classifier, a recent and promising method for pattern recognition, with a fast training algorithm and good accuracy results. Classification performance has been compared in terms of extraction time and accuracy for five moment invariant algorithms, Hu, Legendre, Zernike, Fourier-Mellin and Tchebichef moments, and for three objects different in size, revealing that Tchebichef moments is superior.**

Keywords—Moment Invariants, Optimum-Path Forest (OPF), 3D Sensor, Computer Vision.

I. INTRODUCTION

Object recognition is a complex task in any kind of application of computer vision, because it involves image acquisition, preprocessing, feature extraction and classification [1]. Particularly, in the context of industrial applications, timing requirements play an important role, since classification is supposed to happen in production time. This task cannot therefore, be assigned to humans, because they are usually slower than machines and human experts are hard to find, train and keep in an industry. Moreover, in certain applications, precise information must be quickly or repetitively extracted and used [2], and humans are not the best choice, since they usually get tired quickly.

On the other hand, the raise of high-capacity computers, the availability of high quality and low-priced video cameras, and the increasing need for automatic video analysis has generated an interest in object classification algorithms. A simple classification system consists of a camera fixed high above the interested zone, where images are captured and consequently processed [3].

In this context, this paper describes an application of pattern recognition technique to the object recognition problem based on Optimum Path Forest classifier (OPF) and Moment Invariants approach.

The remainder of the paper is directed as follows. Section 2 describes the feature extraction process. Section 3 gives an overview of classification method. Section 4 discusses the results. Finally, section 5 concludes the paper with a brief discussion of future research.

II. FEATURE EXTRACTION PROCESS

Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which overfits the training sample and generalizes poorly to new samples. Feature extraction is the problem of extracting from the raw data the information which is most relevant for classification purposes, in the sense of minimizing the within-class pattern variability while enhancing the between-class pattern variability.

In object recognition process, the image is to be represented into relevant features, so that the classifier is applied to recognize the object. In this research, Hu, Legendre, Zernike, Fourier-Mellin and Tchebichef moment invariants are used as feature for the object recognition. Based on these methods, each image is represented by a set of features vector.

A. Hu Moments

M. K. Hu [4] introduced a theory of two-dimensional moment invariants for planar geometric figures. This theory introduced seven nonlinear functions which are translation, scale and rotation invariation. This allow compute the center of mass of the image, and of a region in case of a binary mask.

Centralized moment are geometric moments of the image computed relatively to the center of mass. Centralized moments are invariant under translation. To enable invariance to scale, normalized moments are used. Thus, the seven moments are computed.

B. Zernike Moments

F. Zernike [5], introduced a set of complex polynomials which form a complete orthogonal set over the interior of the unit circle, $x^2 + y^2 = 1$

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

where $j = \sqrt{-1}$, *n is a* positive integer or zero, *m* is a positive and negative integers subject to constraints *n*-|*m*| even, $|m| \le n$, 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

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

The discrete form of the Zernike moments of an image size N×N is expressed as follows [6],

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

C. Legendre Moments

The Legendre Moment Invariant was introduced by M. R Teague [7] which is produced based on the recurrence relation of Legendre polynomial of order *p,* is defined as

$$P_p(x) = \frac{(2p-1)xP_{p-1}(x) - (p-1)P_{p-2}(x)}{p} \quad (7)$$

where $P_0(x) = 1$, $P_1(x) = x$ and $p > 1$. Since the region of definition of Legendre polynomial is the interior of [-1, 1], a square image of *N*×*N* pixels with intesity function I_{ij}, $0 \le x, y \le$ (N-1), is scaled in the region $-1 < x, y < 1$.

The discrete form of the Legendre moments of order (*p*+*q*) can be expressed as

$$L_{pq} = \lambda_{pq}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} P_p(x_i)P_q(y_j)I_{ij}, \quad (8)$$

where the normalizing constant is

$$\lambda_{pq} = \frac{(2p+1)(2q+1)}{N^2} \quad (9)$$

x_i an y_j denote the normalized pixel coordinates in the range [-1, 1], which are given by

$$x_i = \frac{2i}{N-1} - 1 \; and \; y_j = \frac{2j}{N-1} - 1 \quad (10)$$

D. Fourier-Mellin Moments

We consider the use of the orthogonal Fourier-Mellin moments introduced by Y. Sheng and L. Shen [8], which are based on a set of radial polynomials, where the radial polynomials an image I_{xy}, $Q_p(x,y)$, are given by

$$Qp(x,y) = \sum_{s=0}^{p}(-1)^p + s\frac{(p+s+1)!\,(x^2+y^2)^{s/2}}{s!\,(p-s)!\,(s+1)!} \quad (11)$$

where *p* is an integer such that $p \ge 0$ and $|q| \ge 0$. In the polar form, $r = \sqrt{x^2 + y^2}$.

The image function is defined over discrete square domain of *N*×*N* pixels and $M^*_{pq}(x,y)$ are the complex conjugate of the complex orthogonal polynomials $M_{pq}(x,y)$, given by

$$M_{pq}(x,y) = Q_p(x,y)e^{jp\theta} \quad (12)$$

where $j = \sqrt{-1}$ and $\theta = \tan^{-1}\left(\frac{y}{x}\right)$, $\theta \in [0, 2\pi]$.

The discrete form of the Fourier-Mellin moments normally used is

$$O_{pq} = \frac{p+1}{\pi}\sum_{i=0}^{N-1}\sum_{k=0}^{N-1} I(x_i, y_k)M^*_{pq}(x_i, y_k)\Delta x_i \Delta y_k \quad (13)$$

$$x_i^2 + y_k^2 \le 1$$

where

$$x_i = \frac{2i+1-N}{D}, \qquad y_k = \frac{2k+1-N}{D}, \quad (14)$$

$$\Delta x_i = \Delta y_k = \frac{2}{D}, \qquad i,k = 0,1,\dots,N-1 \quad (15)$$

and

This work has been supported by CAPES and FUNCAP (PP1-0033-00032.01.00/10).

$$D = \begin{cases} N, & \text{for inscribed circle} \\ N\sqrt{2}, & \text{for outer circle} \end{cases} \quad (16)$$

E. Tchebichef Moments

The discrete Tchebichef polynomials are defined for A. Erdelyi et al. [9] and based on this polynomials, R. Mukunda [10] defined the scaled Tchebichef polynomials as

$$t_p(x) = \frac{(2p-1)t_1(x)t_{p-1}(x) - (p-1)\left(1 - \frac{(p-1)^2}{N^2}\right)t_{p-2}(x)}{p} \quad (17)$$

$p = 2, 3, \ldots, N\text{-}1$

where, $t_0(x) = 1$, $t_1(x) = (2p + 1 - N)/N$.

Under the above transformation, the squared-norm of the scaled polynomials gets modified according to the formula

$$\rho(p, N) = \frac{N\left(1 - \frac{1}{N^2}\right)\left(1 - \frac{2^2}{N^2}\right)\ldots\left(1 - \frac{p^2}{N^2}\right)}{2p + 1} \quad (18)$$

$p = 0, 1, \ldots, N - 1$

Then, the Tchebichef moments are defined as

$$T_{pq} = \frac{1}{\rho(p, N)\rho(q, N)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} t_p(x) t_q(y) I_{xy} \quad (19)$$

$x, y = 0, 1, \ldots, N - 1$

III. Classification

Classification is the final stage of any image-processing system where each unknown pattern is assigned to a category. The degree of difficulty of the classification problem depends on the variability in feature values for objects in the same category, relative to the difference between feature values for objects in different categories M. Mercimek [11]. In this study we use OPF (Optimum-Path forest), classifier as pattern classifier.

A. 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) (A. X. Falcão et al. [13].

Given a training set with samples from distinct classes, the classifier assign the true class label to any new sample, where each sample is represented by a set of features and a distance function measures their dissimilarity in the feature space. The training samples are then interpreted as the nodes of a graph, whose arcs are defined by a given adjacency relation and weighted by the distance function. It is expected that samples from a same class are connected by a path of nearby samples.

Therefore, the degree of connectedness for any given path is measured by a connectivity (path-value) function, which exploits the distances along the path. In supervised learning, the true label of the training samples is known and so it is exploited to identify key samples (prototypes) in each class. Optimum paths are computed from the prototypes to each training sample, such that each prototype becomes the root of an optimum-path tree (OPT) composed by its most strongly connected samples. The labels of these samples are assumed to be the same of their root.

The OPF computes prototypes as the nearest samples from different classes in the training set. For that, a Minimum Spanning Tree (MST) is computed over that set, and the connected samples with different labels are marked as prototypes. A path-cost function that calculates the maximum arc-weight along a path is used, together with a full connectedness graph. Therefore, the training phase of OPF consists, basically, of finding prototypes and execute OPF algorithm to determine the OPTs rooted at them.

Further, the test phase essentially evaluates, for each test sample, which training node offered the optimum-path to it [14].

It has shown good results in many classification problems [15], [16].

IV. 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 real-time applications of classification algorithms. The device has been used to acquire images of three packages with a few differences in size, as shown in Figure 1, 2 and 3.

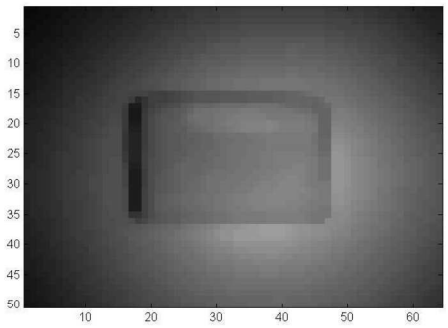

Figure 1. Package 1 with dimension 15 × 10.5 × 7.2 cm.

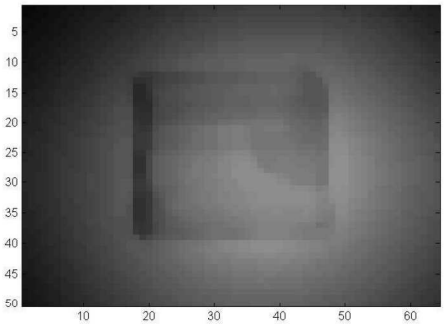

Figure 2. Package 2 with dimension 15 × 14 × 6 cm.

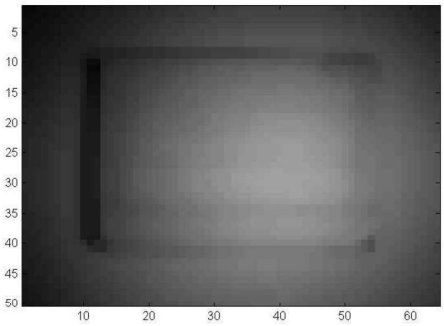

Figure 3. Package 3 with dimension 21.5 × 16.2 × 9.6 cm.

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

In our experiments, the number of input features extracted is shown in Table I.

TABLE I. NUMBER OF INPUT FEATURES EXTRACTED PER MOMENTS.

Moments	Number of Input
Hu	7
Zernike	36
Legendre	36
Fourier-Mellin	36
Tchebichef	36

This numbers of input feature extracted reference to 7 Hu moments and 6 moment of order p and 6 repetitions q, which form a vector with 36 characteristics. The choice of p and q, was made in order to improve the feature extraction time of the moments. In many problems in the image analysis, the order and repetition is random, as may be seen in [6], [7] and [10].

These input vectors are presented to the Optimum-Path Forest classifier. The classifier was trained and tested 20 times with the same database, and the experimental results showed that the recognition rate of the OPF classifier based on Tchebichef moments is higher than the recognition rate of other moments. The results are given in Table II and the Table III shows the processing times, where **Min** - Minimum, **Max** - Maximum, **Mean** - Mean, **Median** - Median and **STD** - Standard Deviation, accuracy.

TABLE II. RECOGNITION RATE OF HU, ZERNIKE, LEGENDRE, FOURIER-MELLIN AND TCHEBICHEF MOMENTS USING OPF.

Moments	Accuracy (%)				
	Min	Max	Mean	Median	STD
Hu	47.5	65.0	57.27	57.5	4.425
Zernike	75.0	80.0	75.75	75.0	1.831
Legendre	72.5	82.5	77.62	77.5	2.747
Fourier-Mellin	70.0	80.0	71.12	70.0	2.747
Tchebichef	82.5	90.0	84.75	82.5	3.526

TABLE III. PROCESSING TIMES.

Moments	Feature Extraction (seconds)	Training (seconds)	Classify (seconds)
Hu	0.555	0.000099	0.000146
Zernike	6.800	0.00012	0.000201
Legendre	0.597	0.000138	0.00015
Fourier-Mellin	7.200	0.000124	0.000147
Tchebichef	0.550	0.000132	0.000152

The Tchebichef moments have better image representation capability than the traditional continuous moments, because Tchebichef moments preserve almost all the information of the image in few coefficients. This is probably assigned to the fact that the orthogonality of methods like, for example, the Legendre polynomials are affected when the image is discretized. As a consequence, such discretization may cause numerical errors in the computed moments.

Feature descriptors that are invariant with respect to rotations in the image plane, can be easily constructed using Zernike moments. Zernike moments demand however more computational efforts than Tchebichef moments. Legendre and Tchebichef moments fall into the same class of orthogonal moments defined in the Cartesian coordinate space, where moment invariants (particularly rotation invariants) are not readily available

In order to evaluate the feasibility of real-time industrial applications, by combining the data from the above tables into Table IV, we can identify Tchebichef and Legendre as more suitable, from the viewpoint of timing requirements of such applications, where **T** - Tchebichef, **L** - Legendre, **Z** - Zernike, **FM** - Fourier-Mellin and **H** - Hu Moments and **S** - total time (seconds).

TABLE IV. ACCURACY × PROCESSING TIMES.

	Accuracy(%)					S
	Min	**Max**	**Mean**	**Median**	**STD**	
T	82.5	90	84.75	82.5	3.526	0,55028
L	72.5	82.5	77.62	77.5	2.747	0,59728
Z	75.0	80.0	75.75	75.0	1.831	6,80032
FM	70.0	80.0	71.12	70.0	2.747	7,20027
H	47.5	65.0	57.27	57.5	4.425	0,55524

On the other hand, the huge running time of Zernike and Fourier-Mellin moments limits their use in this task. This fact can be assigned to the large number of summations in the equations; at the computational level sums are translated into iterative loops, which in turn require much computational efforts.

V. CONCLUSIONS

This paper introduced a comparative study of five moments feature extraction methods (Hu, Zernike, Legendre, Fourier-Mellin and Tchebichef) to the recognize the images of an industrial sensor using Optimum-Path Forest classifier. The experimental results showed that the recognition rate of the OPF classifier based on Tchebichef moments is higher than of the other moments and has a short process time.

Indeed, the accuracy obtained with the best extractor, Tchebichef, achieved 90%. It should be mentioned, however, that 10% of errors is too huge to consider in a real process. This fact can be related to the lack of quality, noise and imperfections on the lighting contained in images extracted from the 3D sensor.

Thus, upcoming activities can be related to the search for an industrial device with better resolution, low noise and controlled lighting, as well as to the test of other feature extraction algorithms, such as Pollaczek, Krawtchouk, Gengenbauer, among others.

ACKNOWLEDGMENT

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

REFERENCES

[1] R. Muralidhran, Dr. C. chandrasekar, (2011). Object Recognition using SVM-KNN based on Geometric Moment Invariant. International Journal of Computer Trends and Technology, July to Aug.

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

[3] P. Kamavisdar, S. Saluja, S. Agrawal, (2013). A Survey on Image Classification Approaches and Techniques. International Journal of Advanced Research in Computer and Communication engineering, Vol. 2, Issue 1, January.

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

[5] F. Zernike (1934), Physica, Vol.1, pp.689.

[6] M. R. Teague, (1980). Image analysis via the general theory of moments. Journal Optical Society American, vol. 70, No. 8, pp. 920-930.

[7] C. W. Chong, P. Raveendran, R. Mukundan (2004). Translation na scale invariants of Legendre moments, Patter Recognition 37, pp. 119 – 129.

[8] Y. Sheng, L. Shen, (1994). Orthogonal Fourier-Mellin moments for invariant pattern recognition, J. Opt. Soc. Am. 11 1748-1757

[9] A. Erdelyi, W. Magunus, F. Oberhettinger, F. G. Tricomi, (1953). Higher Transcendental Functions. New York: McGraw-Hill, Vol. 2.

[10] R. Mukundan, S.H. Ong, P.A. Lee, (2001). Image analysis by Tchebichef moments, IEEE Trans. Image Process. 10 (9) 1357–1364.

[11] M. Mercimek, K. Gulez and T. V. Mumcu, (2005). Real object recognition using moment invariants. Sadhana Vol. 30, Part 6, December, pp. 765–775

[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 19(2), 120-131.

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

[14] M. P. Ponti-Jr and J. P. Papa, (2011). Improving accuracy and speed of optimum-path forest classifier using combination of disjoint training subsets, in 10th Int. Work. on Multiple Classifier Systems (MCS 2011) LNCS 6713 . Naples, Italy: Springer, pp. 237–248

[15] J. Papa, A. Spadotto, A. Falcao, and J. Pereira, (2008). Optimum path forest classifier applied to laryngeal pathology detection, in Systems, Signals and Image Processing. IWSSIP 2008. 15th International Conference on, june 2008, pp. 249 –252.

[16] J. P. Papa, A. X. Falcao, G. M. de Freitas, and A. M. H. de Avila, (2010). Robust Pruning of Training Patterns for Optimum-Path Forest Classification Applied to Satellite-Based Rainfall Occurrence Estimation, IEEE Geoscience and Remote Sensing Letters, vol. 7, pp. 396–400.