A COMPARISON BETWEEN HOUGH TRANSFORM AND MOMENT INVARIANT TO THE CLASSIFICATION OF OBJECTS FROM LOW-RESOLUTION INDUSTRIAL SENSOR IMAGES

R. D. C. SILVA, G. A. P. THÉ

Universidade Federal do Ceará, Depto. de Eng. de Teleinformática, Campus do Pici s/n, Fortaleza-CE

rodrigodalvit@hotmail.com, george.the@ufc.br

Abstract— In this paper, the issue of object recognition, using images extracted from a 3D sensor is discussed. We focus on Hough Transform and moment invariants based feature extraction algorithms for the classification of images from an industrial 3D sensor, using the Optimum-Path Forest (OPF) classifier. 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 Hough Transform, for three objects different in size, revealing which of these is superior.

Keywords: Computer vision and intelligent processing of images, Hough Transform, Moment invariants, Optimum-Patch Forest, 3D sensor.

1 Introduction

Machine vision provides innovative solutions in the context of industrial automation. Indeed, many industrial activities have benefited from the application of machine vision technology on manufacturing processes; a classical example is the size-based separation of objects in conveyor belts of post office companies. Object recognition, however, is a complex task in any kind of application of computer vision, because it involves image acquisition, preprocessing, feature extraction and classification.

Given the variety of commercially available software/hardware solutions for image acquisition and preprocessing, among those tasks feature extraction and classification do appear as the problems to be addressed in image-based control systems. Although high-capacity computers are available, the complexity of the mathematics involved in these two tasks makes image-based control such a non-trivial and challenging problem in industrial environment.

A widely used technique for image analysis are the moments, which are used as descriptors of shapes in a variety of applications, pattern recognition, object classification, face recognition, edge detection, robotic vision, among others. In many of these the extracted features are required to be invariant to scaling, rotation and translation.

Another approach to image analysis is the Hough Transform (HT), which is a powerful distortion and noise tolerant technique that maps image-space points into curves in a parameter or accumulator space.

In this context, this paper describes an application of pattern recognition technique using the HT and Moment Invariants approach to the object recognition problem based on Optimum Path Forest classifier (OPF), a classifier recently proposed (2000s) which has no parameters, and it has shown good results in many classification problems (J. P. Papa et al. (2008)), (J. P. Papa et al. (2010)).

2 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, HT and Hu, Legendre, Zernike, Fourier-Mellin, 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.

2.1 Hu Moments

M. K. Hu (1962), 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. 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. 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 (D. Sridhar, Dr I. V. M. Krishna (2012)).

2.2 Zernike Moments

F. Zernike (1934), 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, $|m| \leq n$, 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 counter-clockwise 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],

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

2.3 Legendre Moments

The Legendre Moment Invariant was introduced by M. R Teague (1980) 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}, $0 \leq x, y \leq (N\text{-}1)$, 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

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

2.4 Fourier-Mellin Moments

We consider the use of the orthogonal Fourier-Mellin moments introduced by Y. Sheng and L. Shen (1994), 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 $p \geq 0$ and $|q| \geq 0$. 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 $\theta = \tan^{-1}\left(\frac{y}{x}\right)$, θ □ [0, 2π].

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

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

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

where

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

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

and

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

2.5 Tchebihef Moments

The discrete Tchebichef polynomials are defined for A. Erdelyi et al. (1953) and based on this polynomials, R. Mukunda et al. (2001) 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

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

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

Then, the Tchebichef moments are defined as

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

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

2.6 Hough Transform (HT)

The Hough transform (HT) (P. Hough, (1959)) is one of the classical computer vision techniques which dates 50 years back. It was initially suggested as a method for line detection in edge maps of images but was then extended to detect general low-parametric objects such as circles. It has been recognized as a very powerful tool for the detection of parametric curves in images (R. O. Duda, P. E. Hart, (1972)); (P. V. C. Hough, (1962)).

The basic functionality of the HT is to detect straight lines. An infinite length os straight line can be described by the equation:

$$y = m * x + b \quad (20)$$

where *(x, y)* are coordinates of points in image space and *(m, b)* are two parameters, the slope and the *y*-intercept respectively.

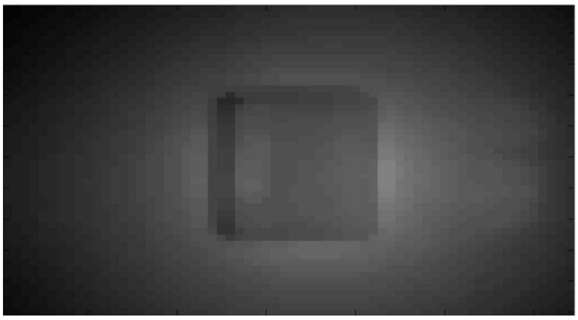
Due to the fact that perpendicular lines to the x-axis can give unbounded values for parameters m and *b* (*b* and *m* rises to infinity), lines are parameterized in terms of theta θ and *r* such that:

$$r = x * cos(\theta) + y * sin(\theta) \quad (21)$$

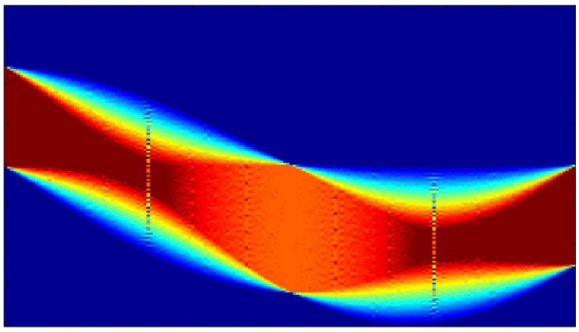
$$\theta \in [0, \pi]$$

where *r* is the length of the vector and θ is the angle it makes with the *x*-axis. Thus, given *x* and *y*, every line passing through point *(x, y)* can uniquely be represented by *(θ, r)*. Both θ and *r* have finite sizes.

The figures below shows an object, box, and its corresponding representation in the Hough space.



(a)



(b)

Figure 1. a) Box detection using Hough transform (original image) b) Hough space correspond to (a).

3 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 (2005). In this study we use OPF, classifier as pattern classifier.

3.1 Optimum-Path Forest (OPF)

J. P. Papa et al. (2009) 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. (2004)).

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 (M. P. Ponti-Jr, J. P. Papa (2011)).

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 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 2, 3 and 4.

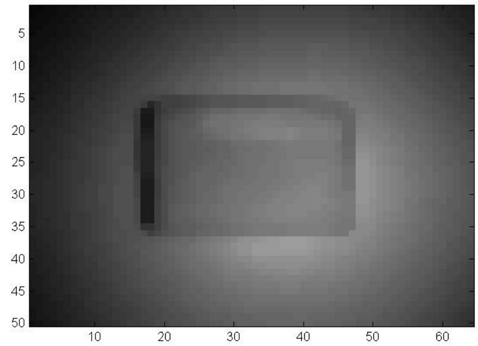


Figure 3. Package 1 with dimension 15×10.5×7.2 cm.

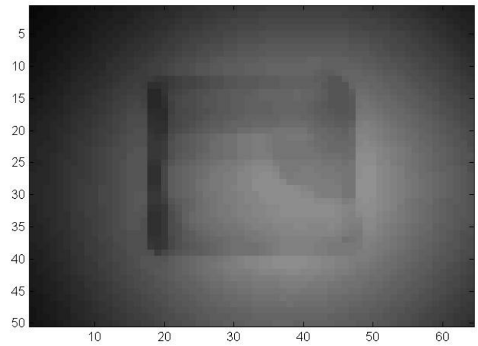


Figure 4. Package 2 with dimension 15×14×6 cm.

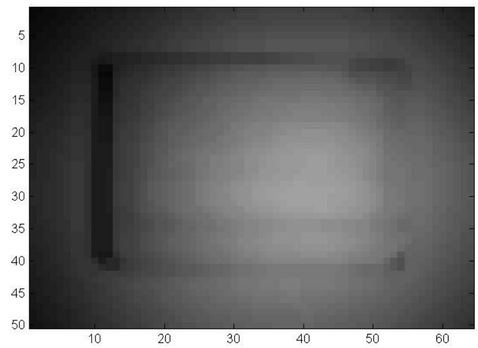


Figure 5. Package 3 with dimension 21.5×16.2×9.6 cm.

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 to the training, contains 18 objects.

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

TABLE I. NUMBER OF INPUT FEATURES EXTRACTED.

	Number of Input
Hu	7
Zernike	36
Legendre	36
Fourier-Mellin	36
Tchebichef	36
Hough Transform	11

This numbers of input feature extracted reference to 7 Hu moments and 0-5 moment of order *p* and 0-5 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 (M. R. Teague, (1980)), (R. Mukunda et al. (2001)).

The Zernike moment, used 11 orders and 11 repetitions. This choice, is due the fact that various moments, such as the moments (p, q) = (0, 1),(0, 2),(0, 3), ..., to be zero, so, it was necessary 11 orders (0 - 10) and 11 repetitions (0 - 10) to form a vector of the same length of the earlier moments, 36 inputs.

The extraction of the characteristic vector of the HT, consists in taking the Hough space features through of the method of Principal Component Analysis (PCA). The PCA (Hotelling, (1933)) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated data into a set of values of uncorrelated variables called principal components, thus. Thus, it was possible to reduce the dimension of the Hough space, (161×180), to a vector with 11 characteristics. Besides the PCA to improve the classification rates of HT, we used a filter Laplacian of Gaussian (LoG) to highlight the edges of objects.

These input vectors are presented to the Optimum-Path Forest classifier. The classifier was trained and tested 20 times with the same database, which contained 150 samples for testing.

The experimental results showed that the recognition rates of the OPF classifier based in Zernike and Tchebichef moments, are higher than the recognition rates of other moments and of the Hough Transform. These results can be seen in Table II, where **Min** - Minimum, **Max** - Maximum, **Mean** - Mean, **Median** - Median and **STD** - Standard Deviation, accuracy, and, **T** - Tchebichef, **L** - Legendre, **Z** - Zernike, **FM** - Fourier-Mellin and **H** - Hu Moments and **HT** – Hough Transform.

TABLE II. RECOGNITION RATE OF MOMENTS AND HOUGH TRANSFORM USING OPF.

	Accuracy (%)				
	Min	**Max**	**Mean**	**Median**	**STD**
H	48.5	72.0	58.700	60.25	6.925
Z	95.5	98.0	96.750	97.00	0.835
L	86.5	91.0	89.375	90.50	1.653
FM	85.5	88.5	87.167	87.50	0.713
T	90.5	92.5	91.900	92.00	0.502
HT	69.0	70.5	69.300	69.00	0.545

As this work focuses on an industrial process in real time, it is important to determine the time required to extract features (for each extractor used) and the time for training the classifier, even if this step occurs only once, and classifying objects. Table III shows the average time of each extractor.

TABLE III. PROCESSING TIMES.

	Feature Extraction (seconds)	**Training (seconds)**	**Classify (seconds)**
H	0.2796	0.000051	0.000094
Z	0.4058	0.00012	0.000404
L	0.3075	0.000094	0.000437
FM	3.6303	0.000094	0.000395
T	0.3068	0.000097	0.000442
HT	0.3201	0.000078	0.000361

The Zernike and Tchebichef moments have better representation capabilities of image than the traditional continuous moments, because these moments preserve almost all the image information in a few coefficients. This is probably assigned to 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 Tchebichef and Zernike moments. Fourier-Mellin moments demand, however, more computational efforts than the others 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, Zernike and Legendre as more suitable, from the viewpoint of timing requirements of such applications, where **S** – total time (seconds).

TABLE IV. ACCURACY × PROCESSING TIMES.

	Accuracy(%)					**S**
	Min	**Max**	**Mean**	**Median**	**STD**	
Z	95.5	98.0	96.750	97.00	0.835	0,4063
T	90.5	92.5	91.900	92.00	0.502	0,3073
L	86.5	91.0	89.375	90.50	1.653	0.3080
FM	85.5	88.5	87.167	87.50	0.713	3.6308
HT	69.0	70.5	69.300	69.00	0.545	0.3205
H	48.5	72.0	58.700	60.25	6.925	0.2797

On the other hand, the huge running time of 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.

The optimal classification rate of the Zernike moment, may be related with the use of the moments of the orders e repetitions 0-10, compared with only the moments of the orders e repetitions 0-5 of the other extractors.

As for the Hough transform, can be said that this had obtained an optimum performance, because this transform is affected with the rotation of the images.

5 Conclusions

This paper introduced a comparative study between Hough Transform and five moments (Hu, Zernike, Legendre, Fourier-Mellin and Tchebichef) with methods of the feature extraction for recognizing the images of an industrial sensor using the Optimum-Path Forest classifier. The experimental results showed that the recognition rate of the OPF classifier based in Zernike and Tchebichef moments, are higher than of the other moments and of the Hough Transform, and have a short process time.

Indeed, the precision obtained with the best extractors, Zernike and Tchebichef, reached 98% and 92.5%, however, 7.5%, is an error rate too huge to consider the recognition of images in real time. This fact may be related to the lack of quality, noise, imperfections on the lighting contained in the images extracted from the 3D sensor and small orders and repetitions used, (0-5) p and q.

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. Besides these, one can find other types of classifiers and compare them mainly using the extraction method of the Hough Transform.

Acknowledgment

Authors knowledge CAPES for financial support. FUNCAP has supported this work (PP1-0033-00032.01.00/10).

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

References

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

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.

D. Sridhar, Dr I. V. M. Krishna. Face Recognition using Tchebichef Moments. International Journal of Information & Network Security, Vol. 1, No. 4, October, pp. 243-254.

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

H. Hotelling, "Analysis of a complex of statistical variables into principal components", Journal of Educational Psychology, 25, 417-441, 1933

J. P. Papa, A. Spadotto, A. X. Falcao, 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.

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.

J. P. Papa, A. X. Falcao, G. M. de Freitas, 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.

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

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

M. P. Ponti-Jr, 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

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

P. V. C. Hough, (1959). Machine analysis of bubble chamber pictures. In *Int. Conf. High Energy Accelerators and Instrumentation.*

P. V. C. Hough, (1962). Methods and means for recognizing complex patterns, U.S. Patent 3; 069; 654.

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

R. O. Duda, P. E. Hart, (1972). Use of the Hough transform to detect lines and curves in pictures, Communications of the ACM 15(1): 11–15.

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